ABSTRACT

Title of dissertation: JET PRODUCTION CROSS SECTION MEASUREMENT IN $\sqrt{s} = 5.02$ TeV PP COLLISIONS

Owen Baron
Masters of Science, 2021

Dissertation directed by: Professor Alice Mignerey
Department of Chemistry and Biochemistry

The study of jets offers insights into the nature of the basic partons composing matter, and enables further studies into the nature of the universe. This analysis presents the double-differential production cross section of radius parameter R = 0.3 proton-proton jets in units of pseudorapidity and transverse momentum at $\sqrt{s} = 5.02$ TeV, measured in the Compact Muon Solenoid detector at the Large Hadron Collider experiment located at CERN. Jets are reconstructed using Particle Flow and the anti-k_T algorithms. The methodology of correcting the detector response through Jet Energy Corrections and Bayesian unfolding is described with a detailed explanation. The results from data recorded in CMS in 2015 are shown and compared with leading-order theory-based simulated PYTHIA8 Monte Carlo events.

JET PRODUCTION CROSS SECTION MEASUREMENT IN $\sqrt{s} = 5.02$ TeV PP COLLISIONS

by

Owen David Cadwalader Baron

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Masters of Science
2021

Advisory Committee:
Professor Alice Mignerey, Chair/Advisor
Professor Nick Hadley
Professor William Walters

Preface

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) allows for the study of proton-proton (pp) and heavy ion (AA) collisions at energies up to $\sqrt{s}$ = 13 teraelectron volts (TeV) for pp and $\sqrt{s_{NN}}$ = 5.44 TeV for AA. These studies allow scientists to test the Standard Model of particle physics, look for new physics outside of it, and gain insight not only into how the universe works now, but also the way it is believed to have existed seconds after its creation in the Big Bang as energy-dense perfect liquid made of nuclear matter. Particle jets are one among the many phenomena observed at the LHC. These jets are collimated sprays of particles resulting from the high-energy collision of protons or heavy ions. Jets are a valuable resource for particle collision study, as they can function either as a probe to measure other physics phenomena or as an observable in their own right. Jets can form in the collision of any nuclear matter, including both proton-proton collisions and heavy ion collisions. As an observable in themselves, jets are considered to be direct substitutes for the quarks and gluons (collectively called "partons") from which they spring, meaning that they carry the characteristics of those partons and the results of their interactions. Studying jets can therefore give insight into the nature of the original partons themselves. In a heavy ion collision, jets can be used to study the properties of the Quark-Gluon Plasma (QGP), a "soup" of high-energy quarks and gluons that arises from the high energy and dense nuclear mass of the heavy ions. These particle jets are expected to arise in the very first stages of the collision and then travel outwards through the QGP before being picked up by the detectors, meaning that the interaction of the jets and the QGP can be measured and studied. In comparison, jet measurements from pp collisions are a useful baseline

against which to compare the jets from heavy ion collisions. Additionally, jets can be used to search for physics beyond the standard model. The existence of super-symmetrical particles can be explored by looking for jet events with significant amounts of missing transverse energy [1]. Jet events can give clues to the search for microscopic black holes produced in the collider, as the evaporation of a black hole is expected to form very similar hadronic jets [2]. As such, it is crucial to be able to accurately describe jets in order to study the properties of partons or to use them as a metric for other analyses and measurements.

While jets are one of the most common phenomena in pp collisions, they are still somewhat rare compared to the overall number of collisions occurring inside the detector. This analysis measures the inclusive production cross section of anti-k_T Particle Flow jets with a radius parameter of R = 0.3 in pp collisions at the Compact Muon Solenoid (CMS) detector, one of several major experiments at the LHC. This measurement can serve as a reference point for further studies on the production of jets in both pp and heavy ion collisions, or the interaction of jets with the QGP. Chapter 1 outlines the basic precepts of the Standard Model and the physics behind particle jets. Chapter 2 introduces the LHC facilities and provides an in-depth overview of the CMS detector and its subsystems. The dataset and Monte Carlo (MC) simulation samples are detailed in Chapter 3, and Chapter 4 describes the reconstruction process of the jets from the readings in the various subdetectors, describing the vertex weighting and the jet reconstruction algorithm. Chapter 5 describes the corrections applied to the data and MC in order to reduce the influence of detector effects. Chapter 6 describes the unfolding process and provides an overview of various inclusive jet measurements confirming the validity of the

analysis. Chapter 7 presents the results and discusses conclusions and further possible areas of research.

Acknowledgments

My profound gratitude goes out to my parents, Eugene Baron and Elizabeth Cadwalader, who encouraged my interest in science at a young age. My advisor, Alice Mignerey, allowed me the ideal environment where I was able to complete my study, and provided me with crucial guidance. Ian Laflotte, my colleague at CMS, offered indispensable help and guidance on this analysis, without which it could not have been completed. In addition, I would like to thank my colleagues Raghav Kunnawalkam Elayavalli, Marguerite Tonjes, Christopher Ferraioli, John Stroglas, Sevil Salur, and all those who helped me on my journey.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 The Standard Model

The Standard Model posits seventeen elementary particles whose interactions create and moderate the three fundamental forces governing interactions in the observed universe: electromagnetic, weak nuclear, and strong nuclear. Elementary particles are defined as those which are not observed to have substructure; in other words, they cannot be broken down into smaller particles. These particles can fall into several classification categories based on how they behave and how they are constructed, which can often overlap. Most broadly, they can be divided into those that form matter and those that carry force. Particles that form matter consist of quarks and leptons and are divided into three generations. In the first generation of matter, these correspond to the up and down quarks and the electron and electron neutrino. These are the particles that make up the elements we are familiar with in every day life: protons, neutrons, and electrons, and also the electron neutrino, which mediates radioactive decay via the weak nuclear force. Further generations contain more massive particles and are mainly found as the products of high-energy particle collisions, either from cosmic rays or particle accelerators. These particles all have corresponding anti-particles with an equal mass and an equal but opposite electric charge. The particles which carry force are called gauge

bosons and correspond to the fundamental forces of the universe. The gluon carries the strong force, which mediates the interactions between partons, with a range of roughly 1 femtometer [7]. The photon mediates the electromagnetic force, which is capable of operating at theoretically infinite distances. The Z and W bosons mediate the weak force, which governs the reactions between leptons and quarks and is involved in the radioactive decay of atoms; in accordance with the name, it operates at distances of roughly one thousandth of a femtometer [7].

The final elementary particle currently known is the Higgs boson, which is a scalar boson. Theoretically, the invariant gauge theory used in the standard model predicted that not only the photon and gluon, but additionally the W and Z bosons, should be massless in order to conserve the symmetry of the electroweak force, which unifies the actions of the electric and weak force into one whole. As the W and Z bosons are in fact quite heavy (about 80 and 90 GeV, respectively), Peter Higgs postulated that a field must exist which is capable of breaking electroweak symmetry [8]. The discovery of a particle which acts like the force carrier of this field was discovered in 2013 by CERN and is considered consistent with the Higgs boson [9]. A summary table of elementary quarks and leptons can be seen in Table 1.1.

Particles can also be grouped into two categories based on their spin properties. Bosons are those particles that follow Bose-Einstein statistical states and have an integer spin. In simple terms, there is no limit to the number of these particles that can exist in the same quantum state at the same time. Examples include the gauge and scalar bosons in Table 1.1, as well as larger composite bosons such as the ^4He atom, or any particle made of a quark-antiquark pair. In contrast, fermions are particles with spin of one half

or multiple halves that follow Fermi-Dirac statistics, meaning that in accordance with the Pauli Exclusion Principle, no more than one can occupy the same quantum state at once. Examples include protons, neutrons, and all leptons.

More complex particles are categorized in several other ways as well. Quarks and gluons carry a color charge, which allows them to occupy states in the same hadron without violating the Pauli exclusion principle. Quarks carry a color charge of either red, blue, or green, and anti-quarks carry opposite charges of anti-red, anti-blue, and anti-green. Gluons are more complicated, carrying one of eight possible color states composed of a combination of color and anticolor. Due to the property of color confinement, any free particle must have an overall net color charge of zero, which can be achieved with either a quark-antiquark pair, or a combination of three quarks or three antiquarks. A particle made of a quark-antiquark pair, such as a pion, is referred to as a meson and has an integer spin of either 0 or 1. A particle made from three quarks or three antiquarks, such as a proton or a neutron, is referred to as a baryon, with a half-integer spin. Both baryons and mesons compose another category called hadrons, which are any particles with a net neutral color charge that are composed of quarks and subject to the strong interaction. Finally, for the sake of simplicity, quarks and gluons are often collectively called partons, referring to the fact that they are the parts that make up the nuclear matter.

1.2 Jet Physics

A particle jet, loosely defined, refers to a cone-shaped spray of highly collimated particles originating from the primary vertex, typically in a high-energy particle accelerator

Table 1.1: Summary of Quarks and Leptons

Quarks			Leptons		
Name	Charge	Spin	Name	Charge	Spin
Up (u)	+ 2/3	1/2	Electron	-1	1/2
Down (d)	- 1/3	1/2	Electron Neutrino	0	1/2
Charm (c)	+ 2/3	1/2	Muon	-1	1/2
Strange (s)	- 2/3	1/2	Muon Neutrino	0	1/2
Top (t)	+ 2/3	1/2	Tau	-1	1/2
Bottom (b)	- 2/3	1/2	Tau Neutrino	0	1/2

collision. The focus of this analysis is on particle jets, composed primarily of hadrons and arising from a hard scattering interaction between the partons composing the colliding protons. Hard scattering is defined as an interaction between particles where the momentum transferred is a large fraction of the total momentum of the system, referred to as a high momentum fraction. The physics of parton scattering is shaped by the forces that govern the partons themselves. As quarks and gluons carry color charge, they cannot exist freely but instead in color-neutral groups. Another relevant property is called asymptotic freedom; according to the non-abelian gauge theory governing the interactions of the strong force, the interactions between partons become stronger as they grow more distant and, conversely, become weaker the closer together they are. There are two main consequences of these laws which allow for the creation of particle jets. First, at close range inside the confines of the nucleus, quarks and gluons act essentially as free particles. Second, in order to maintain color neutrality, quarks and gluons cannot exist on their own as free particles outside of a bound state, but must exist as either a quark-antiquark pair of a color and its anticolor (a meson), or a triplet of three quarks or antiquarks with a different color each (a baryon), held together by gluons. It is therefore impossible to

have a single isolated quark as a stable particle. Two bound partons may be separated up until the energy put into separation meets the energy required to create a new bound pair, at which point said new pair will simply form from the energy. Imagine if, instead of breaking, a stretched rubber band simply divided itself into two rubber bands. In the hard scattering of partons (two quarks, or a quark and a gluon, or two gluons), the opposing momenta of the two particles has enough energy to result in the continued repetition of this process, creating a large number of new particles. These newly-formed particles radiate out from the collision along the path of the original quark pair. In the case of heavy ion collisions, this results in the formation of a quark-gluon plasma (QGP) composed of quarks and gluons at high enough energy and density to lead to a state of deconfinement. In deconfinement conditions, the energy density is high enough to create enough potential interactions between partons to the point where individual quarks and gluons are effectively freed from their color confinement and asymptotic freedom restrictions. Under these circumstances, the partons can move freely within the confines of the created QGP. Fractions of a second after formation of the QGP, the partons lose enough energy for color confinement and asymptotic freedom to become enforced, and the partons in the QGP form into discrete hadrons in a process called hadronization. These hadrons travel through the space between the collision vertex and the detector and are picked up by the various components therein. For a detailed description on how the different particles interact with the detector, refer to Chapter 2.

Chapter 2: The Large Hadron Collider

The Large Hadron Collider (LHC) in Geneva, Switzerland, is the major high-energy experimental facility for the European Organization for Nuclear Research (CERN) and the most high-powered particle accelerator in the world. Construction began in 1994, and the LHC came online in the year 2008. The LHC is laid out in a circle with a circumference of roughly 27 kilometers on the French-Swiss border and is between 50 to 175 meters underground. There are currently four main particle detectors installed in the accelerator complex: ATLAS (A Toroidal LHC Apparatus), ALICE (A Large Ion Collider Experiment), CMS (Compact Muon Solenoid), and LHCb (Large Hadron Collider beauty). This study uses data from CMS, a multi-purpose detector that was constructed for the purpose of exploring the Standard Model of physics, including the highly-publicized search for the Higgs boson [10]. The collider currently runs experiments with pp collisions as well as heavy ion AA collisions, which are either Pb-Pb or Xe-Xe. The results of the collisions of heavy nuclei help researchers understand the evolution of the very early universe and its basic composition. In addition, the collisions can result in the creation of rare particles or events which are ordinarily not observed, and the study of which can give insight into the fundamental workings of the universe.

In order for ions to reach the collision energies needed for experiments, a system

of linear accelerators and synchrotrons is used. A linear accelerator is a long, straight accelerator consisting of conducting chambers with alternating electromagnetic fields. Particles will be pulled forwards by the magnetic field of the chamber in front of them and propelled by the field of the chamber behind them. Superconducting magnets surround the accelerator to keep the ions aligned properly. A synchrotron is a ring-shaped accelerator that uses electromagnetic pulses in order to accelerate the particles. It also employs powerful superconducting magnets in order to bend and focus the beam. For a pp collision at CERN, protons from a source of hydrogen are stripped of electrons as they are sped through the linear accelerator and deposited into the four rings of the Proton Synchotron Booster (PSB). Numerous protons are grouped together into bunches. The bunches then proceed to the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before being injected into the main ring [10]. When matter begins to approach the speed of light, it asymptotically gains mass. The standard way to describe the energy of an ion at ultrarelativistic speeds is to use electron volts (eV). At recent collision energies in the LHC, proton-proton (pp) collisions use beams of protons at 6.5 TeV each, for a total center-of-mass collision energy ($\sqrt{s}$ for pp) of 13 TeV. Heavy ion collisions using lead nuclei (PbPb collisions) are regularly run at a center-of-mass energy ($\sqrt{s_{NN}}$ for heavy ions) of 5 TeV. At different positions on the main ring lie the four main detectors: ATLAS, ALICE, CMS, and LHCb. LHCb is designed to study the matter/anti-matter asymmetry of the universe, while ALICE is configured to be primarily a heavy ion detector. ATLAS and CMS were designed as general-purpose detectors with an additional focus on finding and studying the Higgs boson. The research described in this paper is concerned with the CMS detector.

2.1 The Compact Muon Solenoid

The CMS detector was proposed in 1994 in order to study the electroweak symmetry breaking associated with the Higgs boson [11]. Both before and after accomplishing this task in 2013, CMS has also explored many of the other physical processes occurring in both pp and heavy ion collisions, and looked into the search for physics outside of the standard model. The CMS detector is composed of several distinct inner subdetectors, forming a large cylinder with a radius of roughly 7.5 meters and a length of 21.6 meters. In total, it weighs 12500 tons. A diagram of the detector can be seen in Figure 2.1 [3]. Positioned about 3 meters out from the center is the solenoid magnet, which produces a field of 3.8 Tesla, used to deflect charged particles created in the collisions [3]. Directly around the inner core of the detector lies the silicon tracker, made of thin silicon layers. A charged particle passing through will ionize the silicon, recording its trajectory or "track." The momentum of the particle can then be determined by the curvature of the track in the magnetic field. Outside of the tracker is the electromagnetic calorimeter, or ECAL. The ECAL is composed of lead tungstate crystals, designed to stop electrons and photons and convert their energy into light pulses via scintillation. Outside of the ECAL, but still inside of the solenoid magnet, is the hadron calorimeter, or HCAL, designed to stop charged and neutral hadrons and record their energies by using overlapping layers of brass plates and plastic scintillators. Outside of the solenoid are the muon chambers. These are composed of four layers of drift chambers interspersed with an iron yoke to conduct the magnetic field outside of the solenoid. Muons are particles analogous to electrons and positrons but 200 times heavier, and this considerable mass increase means that they will pass straight

through the central detector without stopping. They are therefore distinguishable as the only particles to penetrate past the solenoid and into the outer drift chambers. In addition, there are two hadron forward, or HF, calorimeters (labeled as very forward calorimeters in Fig. 2.1), which are considered part of the HCAL and are positioned past the ends of the main cylinder and along the beam axis. These are meant to record the particles escaping along the beam axis that are not picked up in the coverage of the other detectors. The Centauro and Strange Object Research (CASTOR) detector is a single unit on only one side of CMS sitting 14 meters from the center, past the HF calorimeters, and is designed to look for strange quark matter. Finally, the Zero Degree Calorimeter (ZDC) sits 140 meters down the beam on either side and records very forward neutral and electromagnetically charged particles, often the result of ultra-peripheral collisions. Further information on the detectors can be found in the CMS Technical Design Report [3].

The CMS detector uses a polar coordinate system seen in Fig. 2.1, where the z-axis is oriented counter-clockwise along the beam path in the LHC circle, with the x-y plane perpendicular to the beam axis. The coordinate ϕ is defined as the angle in the x-y plane as measured from the horizontal x-axis. The coordinate θ is defined as the angle away from the beam line in the z-axis. Instead of θ, the angle of a particle relative to the beam is typically described in terms of momentum in the beam direction. Ideally, the measurement of rapidity, y, would be used for beam-direction momentum, as it is invariant in a relativistic frame and particle multiplicity is nearly even over units of rapidity in the central area of interest. y is defined as

$$y = \frac{1}{2} \ln \left[\left(E + q_L \right) / \left(E - q_L \right) \right], \tag{2.1}$$

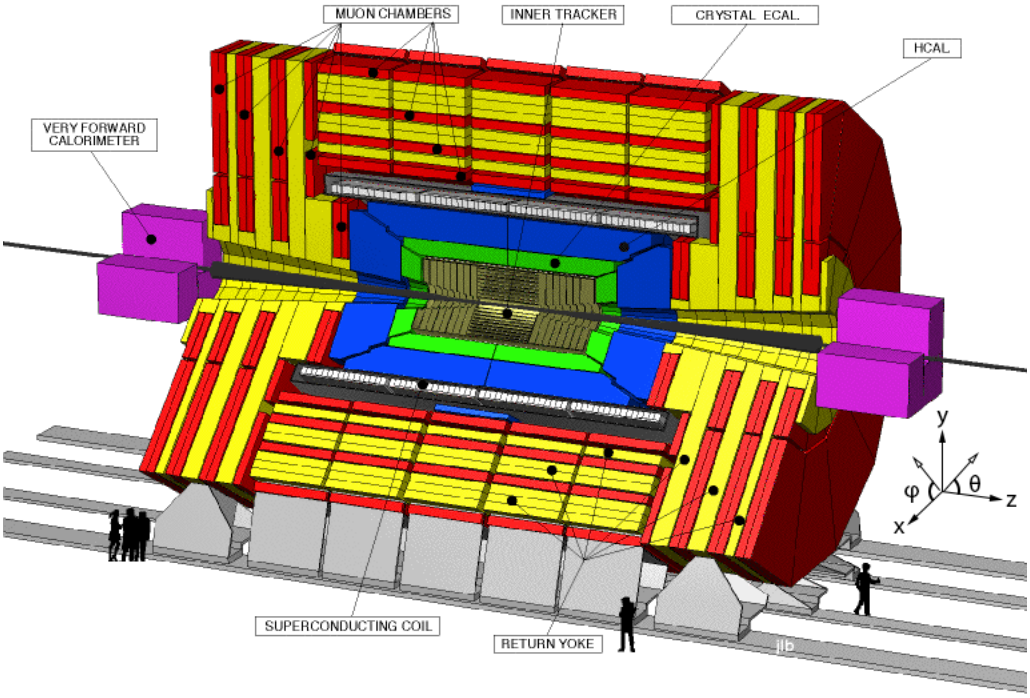

Figure 2.1: Illustrated diagram of CMS detector showing detector components and coordinate system. Adapted from the CMS Technical Design Report [3].

where E is the energy of the particle and q_L is the momentum of the particle along the Z axis [12]. For the purpose of this analysis, pseudorapidity (η) is used as a close equivalent, where:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.2}$$

This construction means that the angles covered by a given unit of η are not constant, and it can be difficult to mentally picture an η-range; however, it is a close approximation of rapidity, allowing an accurate analysis without the need to factor in particle mass and energy identification. In addition, η shows a clear relationship to θ as shown in Fig. 2.2 [13]. Many of the components of the CMS detector have limited coverage in terms of η, which can be seen in Fig. 2.3 [4].

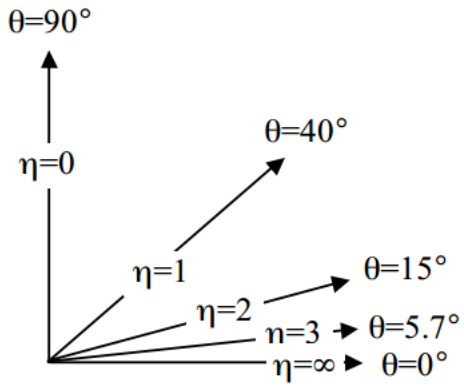

Figure 2.2: Values of η as they relate to the angle θ.

2.1.1 Inner Tracker

The innermost of the CMS subdetectors is the inner tracker, which contains 10 layers of silicon detectors in three major regions. Closest to the interaction point is the pixel tracker, made of silicon pixels measuring 100 x 150 μm in size. The tracker spans a radius from 4 cm to 10 cm in the barrel, consisting of three layers, while each end has two endcaps each at $\pm$35 and $\pm$45 centimeters from the center and radii covering from 6 to 15 cm [11]. The second and third regions compose the silicon strip tracker, sitting outside of the pixel detector. The second or Tracker Inner Barrel (TIB) layer, located 20-55 cm from the beam axis, is composed of four layers of one-sided silicon strips measuring 10 cm x 80 μm. The Tracker Outer Barrel (TOB) region, at 55 cm from the axis, uses silicon strips measuring 25 cm x 180 μm. The endcaps (TEC) are made of 9 levels of microstrip layers, with each layer containing between seven to four concentric circles of tracking strips. Both the barrel regions and the endcaps make use of stereo modules, made of two back-to-back single silicon modules that are rotated at an offset of 100 mrad from

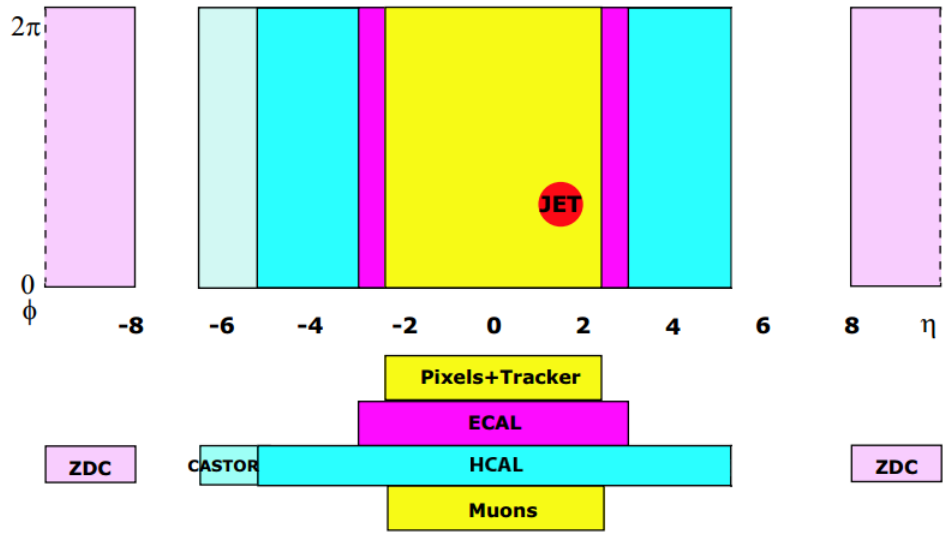

Figure 2.3: η coverage of CMS detector components. Adapted from [4].

each other, to measure particle coordinates in dimensions of r-ϕ and r-z. The silicon strip tracker endcaps sit between $\pm$120 cm and $\pm$280 cm from the central vertex, made of nine disks with seven rings of silicon sensors each, and stereo modules used in the first two and fifth ring. In between the TIB and the endcaps are the Tracker Inner Disks (TID), made of three disks of three rings each, with the first two disks equipped with stereo modules [14, 15].

2.1.2 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (EMCAL) is located outside of the tracking system, and is designed to measure the energy deposited by particles carrying electromagnetic charge. It is constructed of lead-tungstate crystals, which provide a prompt scintillation response and are radiation-hard enough to withstand the intense radioactive environment inside the experiment. Due to the low intensity of scintillation from the crystals, it must

rely on a series of avalanche photodiode detectors in the barrel and vacuum photodiode detectors in the endcap to amplify the signal. Overall, the EMCAL provides coverage over a range of $|\eta| = 3.0$. In addition, the EMCAL contains a pre-shower detector located in front of the endcaps, made of two planes of silicon strip detectors [16].

2.1.3 Hadronic Calorimeter

Outside of the EMCAL sits the main body of the Hadronic Calorimeter (HCAL) subsystem, designed to measure the energy deposited by hadronic particles. The HCAL is a sampling detector made of alternating layers of scintillating plastic and brass. Hadronic particles entering the detector will interact with the brass absorption plates, creating a particle shower. The particles in the resulting shower trigger scintillation in the plastic layers, the light from which is transformed by wavelength-shifting fibers to the visible range and carried by clear optical cables to hybrid photodiodes. The HCAL consists of the Inner (HB) and Outer (HO) barrels, the Hadron Endcaps (HE), and the Hadron Forward (HF); the HB and HE are located inside the detector's magnetic coil, and the HO and HF sit outside [17]. The Inner Hadron Barrel is made of two half-barrels, covering a range of $|\eta| < 1.4$. Inside, 15 brass plates of 5 cm thickness alternate with 17 scintillation layers, of which the first is the thickest at 9 mm and the rest are 3.7 mm. The inner and outermost absorber layers of the HB are made of stainless steel. The Hadron Endcaps (HE) are constructed of alternating layers of brass absorbers and plastic scintillators, divided into 14 towers in each endcap and covering an η range of 1.3 to 3.0. These endcaps allow the HCAL system to detect particles falling outside of the range of the barrel. The HO

sits outside of the magnetic coil in order to catch the tail end of particle showers in the HCAL barrel, designed to reduce the missing energy background in events and improve the missing energy resolution, aiding in the search for non-interacting particles. The barrel of the HO is divided into 5 rings, covering a range of $|\eta| < 1.26$ with several spaces in coverage for structural support features. The central ring has two 10 mm layers of scintillators separated by a 18 cm iron absorber, while the other four rings contain only one scintillation layer. The HF consists of two units on either side of the main detector, sitting 11.2 meters away from the event vertex, following the beam path. These forward detectors allow for an extension of $|\eta|$ coverage from 3.0 to 5.0.

The Zero Degree Calorimeter (ZDC), designed to measure spectator neutrons and forward photons in pp and heavy ion collisions, sits roughly ± 140 m down the beam path from the interaction vertex. It assists in determining luminosity and centrality in heavy ion collisions. The detector is split into hadronic (HAD) and electromagnetic (EM) sections, each composed of alternating layers of tungsten plates and 0.7 mm quartz fibers. The tungsten plates are 15.5 mm and 2 mm thick in the HAD and EM sections, respectively. There are 24 alternating layers of plates and fibers in the HAD, and 33 layers in the EM [18].

2.1.4 Magnet

The 12.9 meter long solenoid magnet at the center of the CMS detector lies between the Hadron Barrel and the Hadron Outer, separating the HO and the Muon Chambers from the inner parts of the detector. The magnet along with the return yoke weighs a total

of 12,500 tons and carries a magnetic field of 3.8 Tesla. The purpose of the magnetic field is to bend charged particles in order to provide a measurement of their momenta. The momentum of a particle can be measured by relating the p_T to the radius curvature, magnetic field, and charge of the particle. The structure of the iron return yoke sits outside the magnet and consists of barrel and endcap structural portions [19].

2.1.5 Muon Chambers

The CMS detector is built with a focus on muon measurement, as the name implies. With the same charge but 200 times more mass than an electron, muons are difficult to detect in smaller detectors as they can penetrate through several meters of solid iron without interacting. The muon detector system in CMS is built into the iron return yoke of the solenoid magnet, outside of the magnet coil and the Hadron Outer, in the range where other particles can be assumed to have already been absorbed by the inner detectors. Muons are capable of being measured in three parts of the detector: the inner tracker, the hadron calorimeter, and the muon chambers. Due to the penetrating nature of the muons, however, they are unlikely to deposit much energy in either the tracker or the HCAL. Three types of detectors are used in the muon chambers: drift tubes (used in the barrel region), cathode strip chambers (used in the endcaps), and resistive plate chambers (used in both barrel and endcap). In the barrel region, four concentric layers of muon Drift Tubes (DT) are interspersed with the iron plates of the return yoke, which additionally serve as absorbers for the muons. This barrel region covers a range of $|\eta| < 1.2$. Each of the three inner layers contains 12 chambers, made of three groups of four drift tube

layers each. Two of these groups measure the r-ϕ coordinates, and the third measures the z coordinate; the fourth layer contains only the groups measuring the r-ϕ coordinates. The drift tubes are coupled with Resistive Plate Chambers (RPC) on either side in the first two barrel layers, and one on the inner side of each of the outer two layers. Cathode Strip Chambers (CSC) are used in the endcaps, where there is an uneven magnetic field and the muon and background rates are higher than in the barrel. The endcaps allow coverage to be expanded to $|\eta| < 2.4$. The endcaps consist of four layers of CSCs and RPCs. The CSCs are trapezoidal in shape and are overlapped in phi to eliminate gaps in coverage. Each chamber consists of a perpendicular series of anode wires and copper cathode strips, which provide measurements in two coordinates. There are RPCs located on the outside of four of the CSC layers, extending up to $|\eta| < 2.1$. Overall, these subsystems combine to give CMS muon detection up to $|\eta| < 2.4$ [20].

Chapter 3: Dataset and Simulation

3.1 Triggers

The probability of interesting physics events occurring in a given collision is very low. In order to ensure enough interesting events are recorded, the LHC takes the approach of producing a very high number of total collisions to ensure as many interesting ones as possible. The rate of collisions at the LHC is on the order of 10^9 interactions per second, but the detector system is only capable of writing the information of around 100 events per second to disk. In order to obtain the most helpful data, the vast majority of the events must be weeded out so that only the ones of interest are recorded. The CMS detector uses a system of triggers to narrow down the events written to disk, consisting mainly of the Level 1 (L1) trigger and the High Level Trigger (HLT). The L1 trigger is at the forefront of event rejection, and therefore must make the choice to accept or reject an event in several fractions of a second. By the end of the L1 trigger process, the rate of events has been narrowed down to around 100 kHz. This is accomplished by using data from the calorimeter and muon chambers, via the use of custom hardware connected to the subdetectors themselves. Each detector uses its hardware to run its own L1 trigger. In order to hasten the acceptance or rejection decision, the L1 trigger is preceded by the Trigger Primitive Generator (TPG), which greatly reduces the resolution

of the information contained in the event [11, 21].

The HLT system is contained on a dedicated server farm, and due to the lower event rate is able to work with information at roughly the same resolution as the offline reconstruction software, allowing the use of reconstruction algorithms nearly identical to their offline counterparts. The HLT uses tracking, vertex reconstruction, Particle Flow jet reconstruction, muon identification, and b-tagging in order to make decisions on which events to accept and reject; the output rate at the end of the process is roughly 800 Hz [11, 22]. The remaining events are piped into storage on the CERN servers.

All data used in this analysis is collected from 2015 runs of proton-proton collisions in the LHC at CERN. The two datasets, HighPtJet80 and HighPtLowerJets, consist of jet events that fire one or more of three triggers resulting in the event being written to the respective dataset. In the case of HighPtLowerJets, an event containing at least a single jet with $p_T > 40$ GeV or $p_T > 60$ GeV is required; for the HighPtJet80 dataset, a single jet with a $p_T > 80$ GeV is required. These criteria are selected for by the use of three triggers, HLT40, HLT60, and HLT80, selecting for any jet event that has a p_T falling within its threshold. Under this system, it is possible that a single event may fire more than one trigger and thus be written to both datasets, so care must be taken when combining them in order to exclude duplicate events. Due to the increasing rarity of jet production as p_T increases, the number of softer, low-p_T jet events recorded is far greater than the the number of high-p_T jet events. If all the events passing at least one of the triggers were simply to be recorded in proportion to their natural occurrence, the resulting sample would be inundated with low-p_T jets and sparse in high-p_T jets. A prescale weight is therefore applied to the events correlating with their triggers in

Table 3.1: Event Selection for Combined HighPtLowerJets and HighPtJet80 Data

Event Selection	Combined Datasets	% of Total
No Selection	$6.04652X10^{07}$	100%
Skim Cuts	$6.03522X10^{07}$	99.8131%
Vertex Cut	$6.03471X10^{07}$	99.8047%

Table 3.2: Event Selection for PYTHIA8

Event Selection	PYTHIA8	% of Total
No Selection	$1.36996X10^{07}$	100%
Skim Cuts	$1.36832X10^{07}$	99.8806%
Vertex Cut	$1.36795X10^{07}$	99.8536%

order to ensure an enhanced selection of high-p_T jets, meaning that relative to their natural production frequency, a lower percentage of jet events with low p_T are recorded, a medium percent of jets with medium p_T are recorded, and a high percent of jets with high p_T are recorded. In the case of jets that fire more than one trigger, a jet with an ambiguous trigger classification will be assigned the prescale weight corresponding to the highest fired trigger. In order to prevent events written to both the datasets from appearing as duplicates, only jets with a prescale weight of HLT80 are used from HighPtJet80, and only jets with a prescale weight of HLT60 or HLT40 are used from HighPtLowerJets. The statistics for the dataset and MC simulation can be seen in Tables 3.1 and 3.2.

3.2 Trigger Efficiency

The online HLT80 and HLT60 triggers are measured for efficiency using the HLT40 and HLT60 triggers as respective reference points. A ratio of the number of events at a given p_T which fire the HLT60 trigger to the number of events at the same p_T firing

the HLT40 trigger can be seen in Fig. 3.1. The blue curve shows a polynomial fit to the entries, used to calculate the lower p_T threshold for the trigger. The same graph for HLT80 efficiency using HLT60 as a reference can be seen in Fig. 3.2. Through this method, the HLT60 trigger is determined to have a minimum p_T threshold of 90 GeV, and the HLT80 value to have a minimum p_T threshold of 110 GeV. When combining datasets, these thresholds are used to ensure only the jets in regions where the trigger is most efficient are counted in the analysis. Plots of the number of events with a given lead jet p_T per each HLT trigger before and after running jet combination are shown in Fig. 3.3 and Fig. 3.4, respectively.

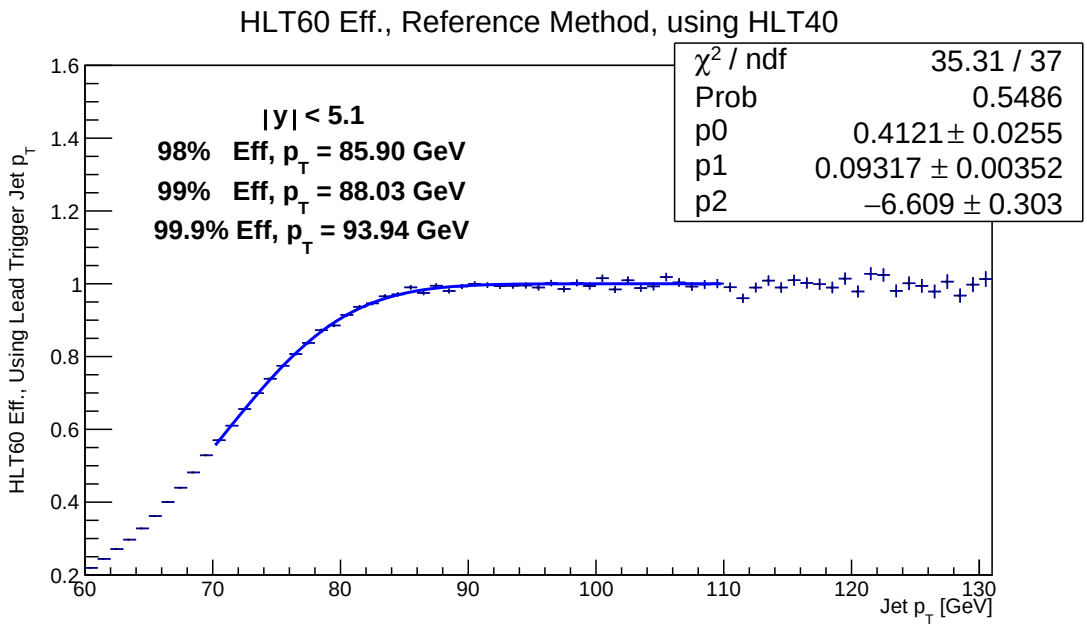

Figure 3.1: Number of events at a given p_T firing the HLT60 trigger plotted as a ratio to the number of the same events at that p_T firing the HLT40 trigger. The blue line represents a polynomial function fit to the ratio.

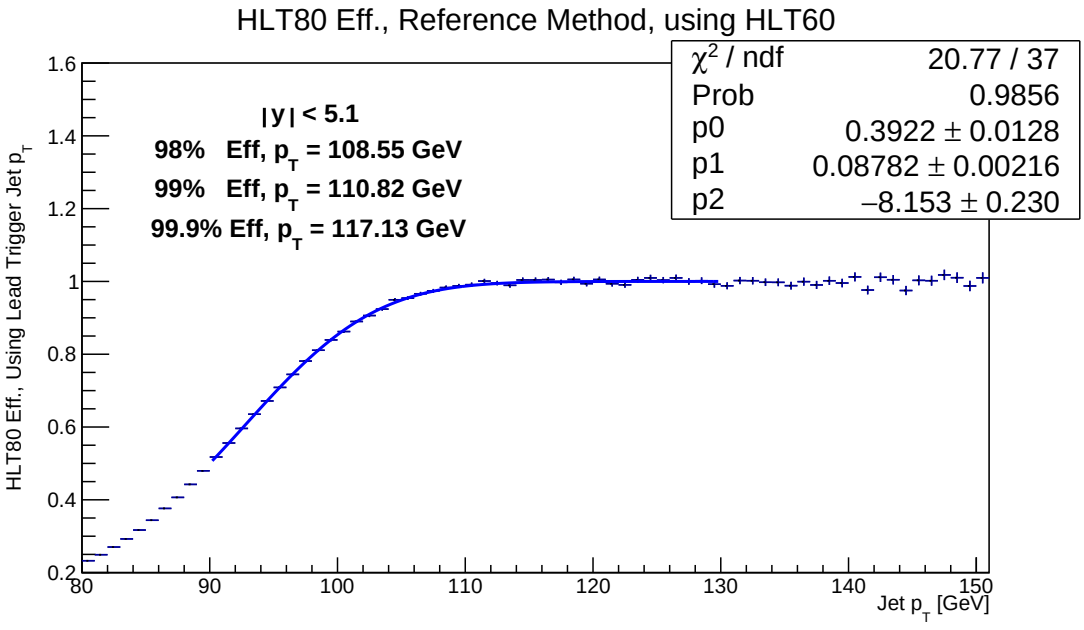

Figure 3.2: Number of events at a given p_T firing the HLT80 trigger plotted as a ratio to the number of the same events at that p_T firing the HLT60 trigger. The blue line represents a polynomial function fit to the ratio.

3.3 Event Selection

In order to increase the quality of the analysis, events are selected for properties relevant to this study. Events must have at least one jet event with a primary vertex within 24 cm in any direction from the detector interaction point, within $|\eta| < 5.1$, and a p_T higher than the relevant trigger of either HLT40, HTL60, or HLT80.

3.4 Leading Order Monte Carlo Simulations

Calculation-based simulation of proton-proton collisions is used to provide a theory comparison to the data collected by the CMS detector, as well as a guide for unfolding and energy corrections. Monte Carlo techniques involve random sampling of a Probability Density Function (PDF) in order to create simulated pp event interactions. This analysis

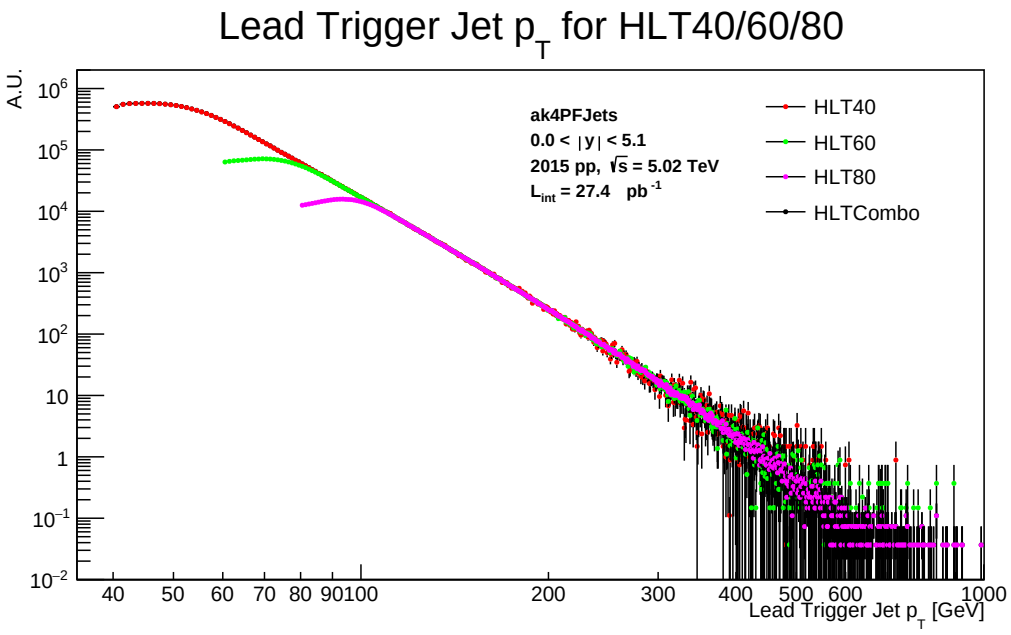

Figure 3.3: Number of events with a given lead jet p_T by HLT trigger, before jet combination code is run. Colors correspond to the three HLTs, and the combined dataset is shown in black.

uses PYTHIA8 dijet simulation using the CUET8PM1 tune [23]. The PDF used is created from Leading Order (LO) calculations of perturbative QCD (pQCD) theory. The events created by the Monte Carlo event generator are known as the Generator Level, GEN-level, or Truth Monte Carlo events, and they represent the particle-level reality of the simulated events. The GEN-level events are fed into a simulated CMS detector constructed in the GEANT4 software platform, which models the interaction of the particles with the matter of the detector in order to create a simulated detector response [24]. The resulting events are recorded in the simulated detector and categorized with the jet-finding algorithm to create the RAW-level MC simulation. Finally, the RAW MC events are corrected using MC-Truth Jet Energy Corrections (JEC) as detailed in chapter 5.1.2 to create the reconstructed, or RECO-level, Monte Carlo event simulation.

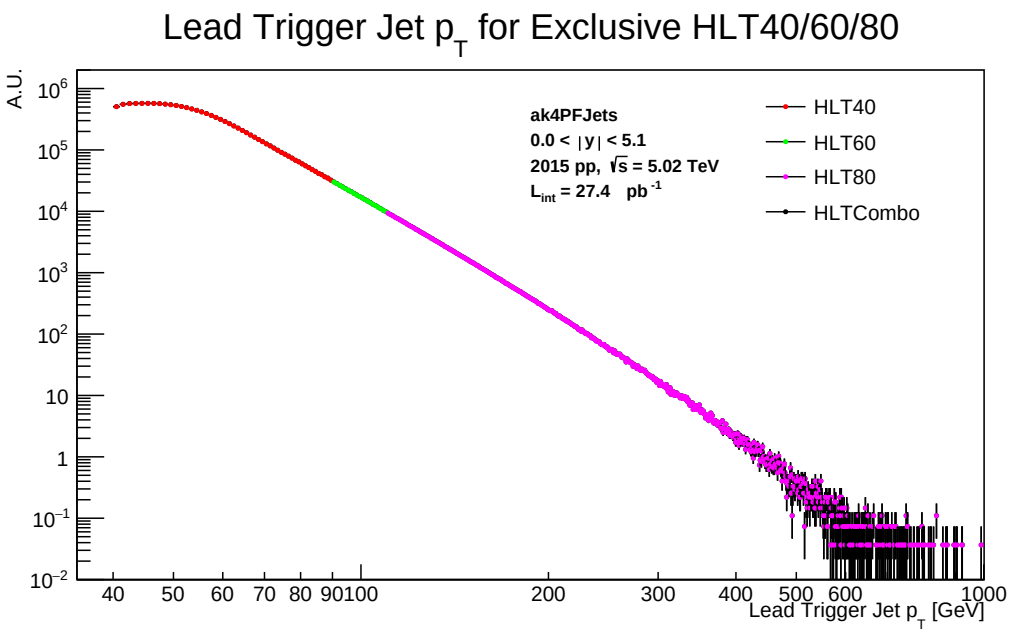

Figure 3.4: Number of events with a given lead jet p_T by HLT, after jet combination code is run, showing only the jets falling within the most efficient pT threshold for the given triggers, corresponding to the colored labels. The combined dataset is shown in black for comparison.

3.5 Event Weights

Eleven different samples of Monte Carlo simulation are produced in the PY8 MC production process to model each of the $\hat{p}_T$ bins used in the analysis. The unweighted number of events per $\hat{p}_T$ bin is shown in Fig. 3.5. In order to create a simulation with increased statistical robustness, high numbers of events are produced for each $\hat{p}_T$ bin, creating on the order of the same number of events for each. In order to make the simulation reflect the reality that events become rarer with increasing p_T, the events in each $\hat{p}_T$ bin are weighted in order to create a continuous spectrum, as seen in Fig. 3.6. As opposed to the uneven function seen in the unweighted spectrum, the weighted spectrum is smooth and continuous, indicating that the weights have been correctly calculated.

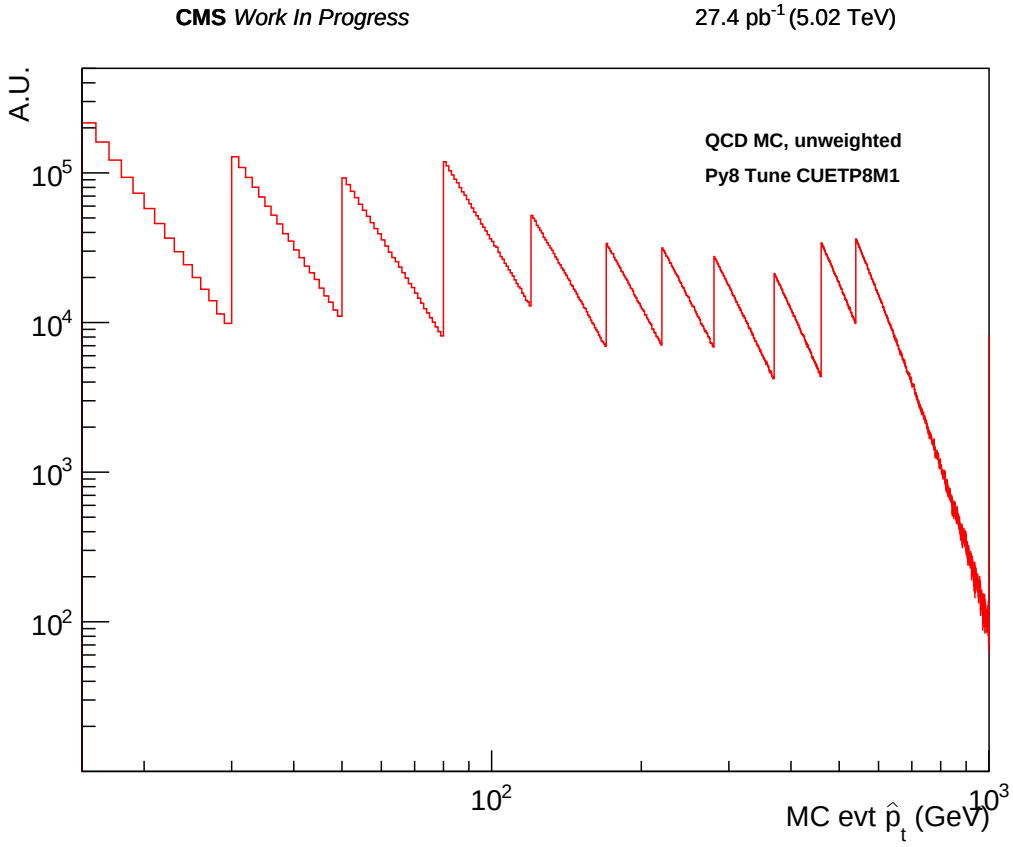

Figure 3.5: Unweighted spectrum for PY8 Monte Carlo simulation.

3.6 Vertex Reweighting

The distance of events from the center collision point of the detector forms Gaussian distributions for both the Monte Carlo simulation and the data collected in the detector, but the distributions are slightly different. As seen in Fig. 3.7, the Monte Carlo event distribution is wider and shorter compared to the distribution of events in data, in part because of the difference between the real detector geometry and the idealized simulated detector. In order to make the distribution of the MC match the data, the event vertexes are reweighted. The ratio of MC to data is calculated and fit with a polynomial from which a weighting factor can be derived. This weight is used to make the distribution of

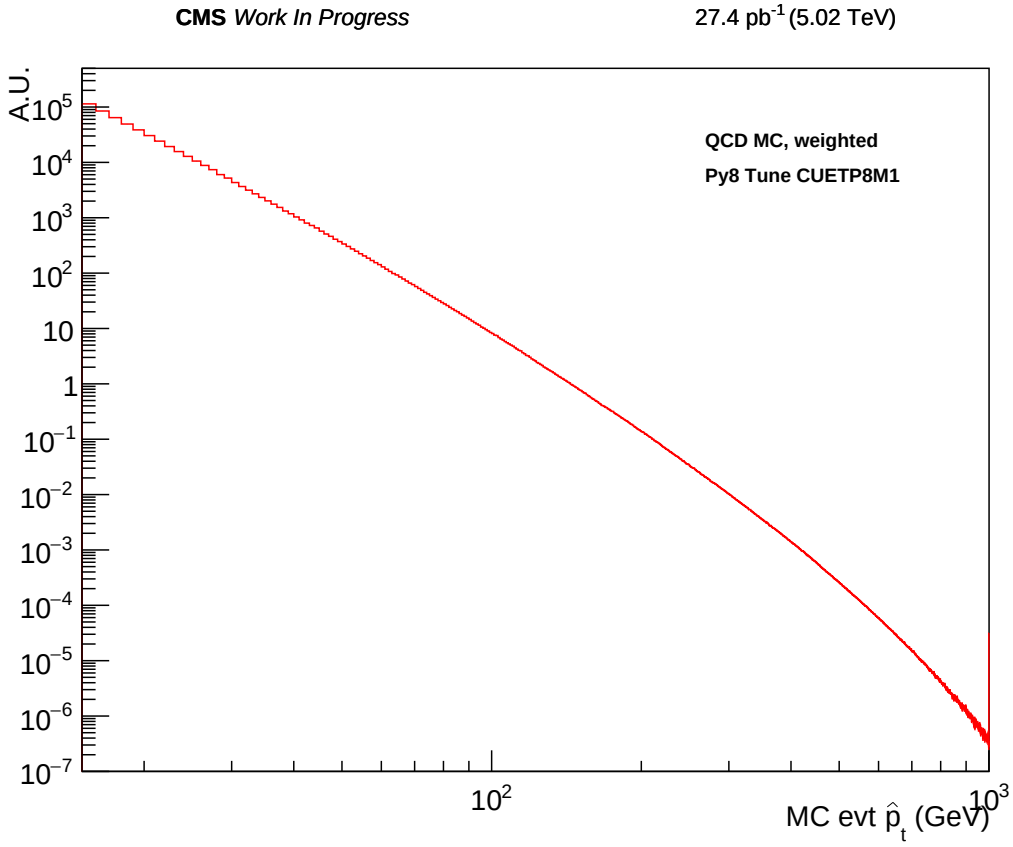

Figure 3.6: Weighted spectrum for PY8 Monte Carlo simulation.

the MC events match that of data, as seen in Fig. 3.8. The ratio of the MC to data is now consistent with unity, as shown in the lower portion of the plot, meaning the weighted MC can be directly compared to the data.

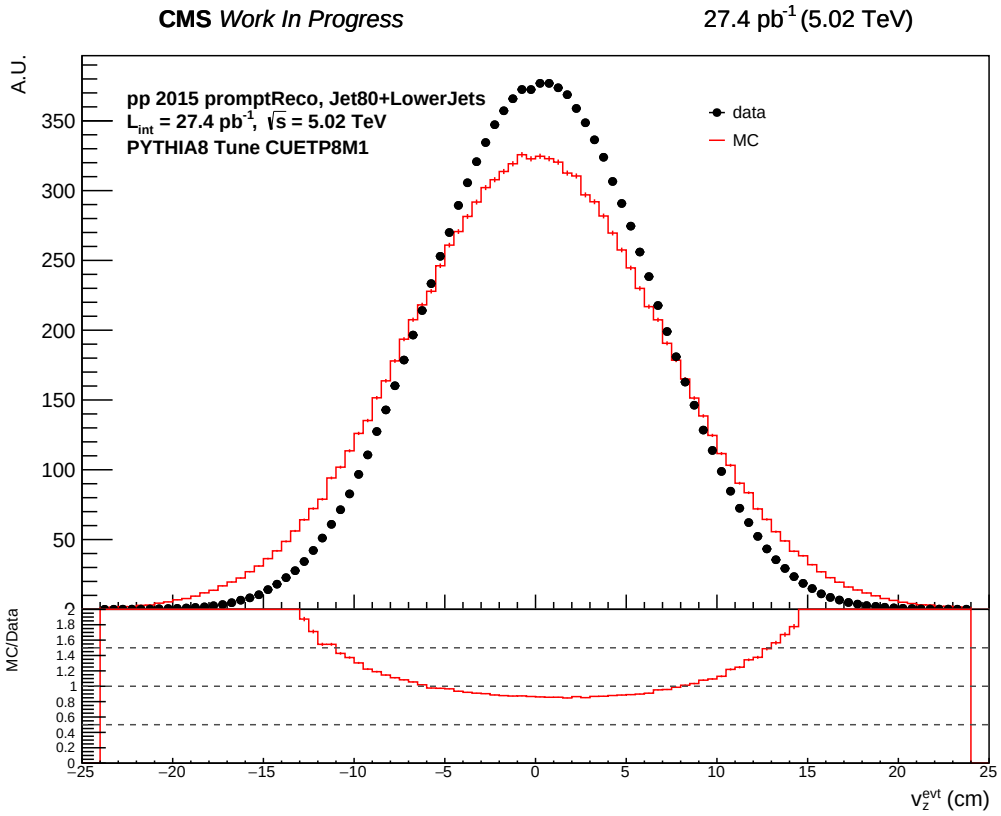

Figure 3.7: Compared unweighted distribution of event vertices from the central interaction point in the CMS detector. The bottom section of the plot shows the ratio of MC to data.

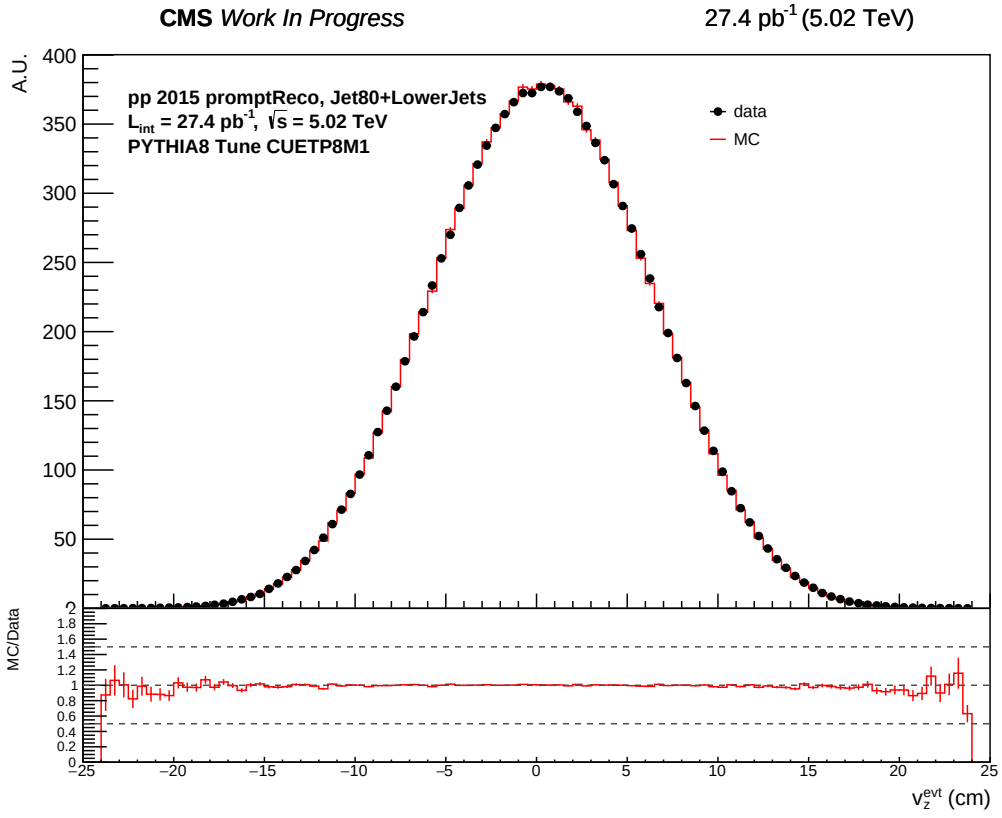

Figure 3.8: Compared weighted distribution of event vertices from the central interaction point in the CMS detector. A ratio centered at unity indicates successful weighting.

Chapter 4: Reconstruction

4.1 Vertex Reconstruction

Collisions in the CMS experiment do not all occur at the exact center of the collider, and so care must be taken to account for events that are shifted relative to the detector geometry. In order to do so, algorithms are used to trace the tracks and deposits in the detector back to the the coordinates where the collision took place, known as the event vertex. The signal from the event also must be separated from noise caused by other incident background particles. The algorithms used at CMS are based on a Kalman filter process in order to find and isolate only the tracks originating from the event vertex in order to eliminate noise from the event [25, 26].

4.2 The CMS Particle Flow Algorithm

In order to property reconstruct a comprehensive view of a jet event from its constituent parts, the algorithms must know what those parts are. In early experiments, only deposits in the ECAL and HCAL were used to determine the structure of jets, which did not allow for identification of all particles. The Particle Flow (PF) algorithm makes use of all the CMS subdetector systems, including the silicon tracker and muon detectors, in

order to reconstruct events with complete particle ID [27]. The Particle Flow algorithm accomplishes this by spatially matching tracks to the presence or absence of corresponding detector clusters. Muons, handled first in the identification process, will leave a track in the silicon tracker and tracks in the outer muon detectors. Charged hadrons leave a track corresponding to deposits in the HCAL and ECAL, where a neutral hadron may leave ECAL deposits as well as HCAL deposits, but without a track. A photon will leave only a deposit in the ECAL. An electron will leave a track corresponding to an ECAL deposit. The strong magnetic field in the CMS detector helps separate the neutral particles from the charged particles, and the curved path of the charged particles can be used to derive their momentum [27, 28]. Via this method, all particles and their basic properties can be identified, then used to reconstruct the jets in a given event.

4.3 Jet Algorithm

A jet-finding algorithm must be used in order to identify jets for study. Jet algorithms traditionally fall into two categories: either cone algorithms or sequential clustering algorithms. Cone algorithms, such as SISCone, construct jets by finding a high-energy object in a jet and extending a radius around it, then summing up the other objects in that area into a trial-jet, the axis of which is then compared to the original seed axis. Iterative clustering algorithms, such as the Cambridge/Aachen, k_T, or anti-k_T algorithms, determine the distance between objects and clusters until no objects are left. This analysis makes use of FastJet software's implementation of the anti-k_T algorithm [5]. The anti-k_T algorithm possesses several advantages compared to other similar algorithms, but is mainly distinguished

for its ability to create a more rounded jet area using a clustering process, which in other algorithms often produces irregularly-shaped jets. A comparison of several different jet-finding algorithms and their areas is shown in Fig. 4.1. The anti-k_T algorithm compares the distance between two detected particles to the distance to the beam direction in order to reconstruct a jet event from clustering the individual particles detected. The formula to compare two objects i and j is based on the distance between two objects, d_{ij}, defined as:

$$d_{ij} = min(k_{ti}^{-2}, k_{tj}^{-2})\Delta R_{ij}^2/R^2 \tag{4.1}$$

where

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{4.2}$$

and also on the distance between the first object and the beam direction, d_{iB}, defined as:

$$d_{iB} = k_{ti}^2. \tag{4.3}$$

Should d_{ij} be the smaller of the two values, the items are recombined into the same jet. However, if d_{iB} is larger, it means the object i is a new jet. This procedure is repeated until all objects have been classified. The radius parameter of this algorithm, R, is traditionally set in the range between 0.4 to 0.7, but this analysis looks at jets with a radius parameter of 0.3. A larger radius includes more of the outer jet substructure but is susceptible to background contamination, while a smaller radius is less contaminated but may ignore the outer parts of the jet in favor of concentrating on the hardest objects near the center. Thus, this analysis will mainly be describing the activity of the inner, harder core of the

jet events, leaving out the softer radiation on the outside edges. As this analysis uses the Anti-k_T jet finding algorithm with a parameter of R = 0.3 using candidates from the Particle Flow algorithm, the jets can be referred to as AK3PF jets as a shorthand.

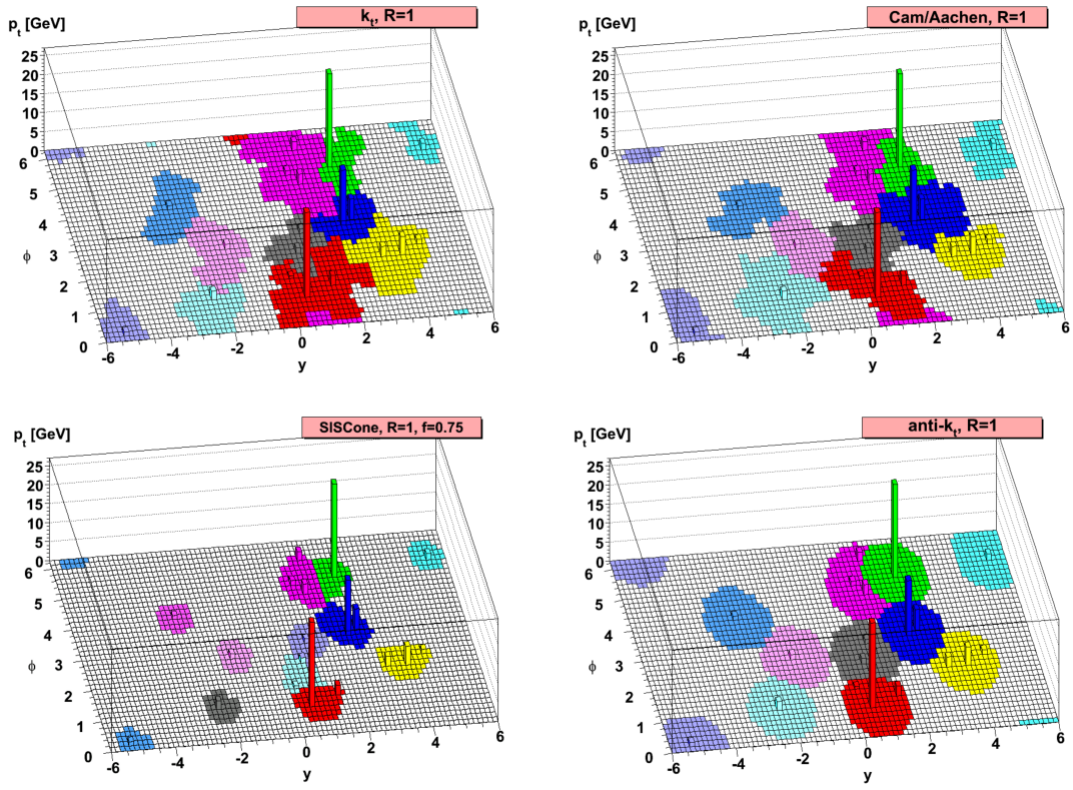

Figure 4.1: A comparison of different jet algorithms given the same events, taken from [5].

Chapter 5: Data Analysis

5.1 Jet Energy Corrections

Several factors in the detection of jets require corrections of the detector response in order to obtain an accurate picture of the actual physics. These can be broken down into several different levels, and in the CMS system corrections for these factors are referred to as the Jet Energy Correction (JEC). The JEC is applied in a sequential factorized basis; corrections are sequentially applied to the jets after they have undergone the corrections from a previous step. The jets must be corrected for pileup offset, then relative and absolute detector response in terms of both p_T and η. For studies focusing on specific flavors of jets, optional corrections can be applied to correct for differences in how the detector measures jets originating from different flavor partons.

5.1.1 Pileup Events

The first applied corrrections (level one, or L1) are to correct for the pileup offset caused by unrelated particles in the background of the event of interest. Due to the nature of collisions at the LHC, multiple bunches of protons are sent at the same time for every beam crossing in order to maximize the potential of an interaction. The beam crossings

happen within fractions of a second, and often the final-state particles of the previous event are still present during the subsequent event. This is referred to as "out-of-time pileup", and can be minimized using calorimeter signal processing [29]. Since there are multiple protons sent out in each beam crossing, the possibility of more than one event occurring in the same bunch exists, leading to extra tracks and energy deposits uncorrelated to the event of interest, known as "in-time pileup". In general, a higher luminosity correlates to higher pileup. Due to the low luminosity of the sample used for this analysis, the JEC applied to the data does not require pileup corrections.

5.1.2 Monte Carlo Simulation-Based Corrections

Level 2 and 3 (L2Relative and L3Absolute) corrections are carried out in one step. The response of the detector is non-linear with regard to increasing transverse momentum and with regard to coverage in η, and thus the response must be scaled back down to match the true energy of the original events. In order to accomplish this, it is necessary to know the extent to which the detector energy is offset. The correction for relative detector response in terms of η is carried out with a process of creating simulated events using the PYTHIA8 and GEANT4 software in order to create corresponding sets of generator and detector-level events, similar to the one carried out to create the PY8 MC simulation for comparison with the data; this process is referred to as MC-Truth correction. As the detector-level jets have not had JEC applied yet, they are referred to as RAW jets. The RAW jets are matched to the GEN jets within the $\eta - \phi$ phase space. The central region of the detector ($\eta < 1.3$) is the easiest to calibrate and has the lowest variance, and therefore

events here are known to the best resolution. The ratio of the GEN jets to the RAW jets in terms of η is used as a correction factor applied to the RAW data to flatten the response. The correction for absolute detector response corrects for the uneven detector response in terms of varying p_T. MC-Truth corrections derived from the same simulation used for the relative response are used to correct the absolute response. Using Jets from the central $\eta < |1.3|$ region of the barrel, the average difference in p_T between GEN jets and spatially-matched RAW jets is measured and used as a ratio in the same manner as the relative correction to create a map to correct the detector response. After this correction, the response is expected to be flat over all bins of p_T.

5.1.3 Residuals

Although the MC-Truth method will provide an adequate correction, several distortions in η and p_T still remain due to the inherent difference between data and simulation. Therefore, corrections derived from data are used in order to make up the difference. The relative jet response in η is corrected for by using the dijet balancing method [30, 31, 32]. In order to correct for other η regions, we examine back-to-back dijets with one jet in the central $\eta < |1.3|$ region and the other jet outside of the central region; these are respectively called the barrel and probe jets. The jets must have a separation of $\Delta\phi > 2.7$ radians to ensure an effective correction. The average p_T in a given η region and the average p_T in the central η region are used as a ratio to create a map in order to correct and flatten the p_T response of jets in outer η regions relative to the central barrel region. The residual data-based correction is derived in the same fashion as the MC-Truth correction,

with the response ratio taken as a weight applied to the jets after MC-Truth corrections have been applied in order to obtain the corrected results. Absolute corrections for p_T response are carried out using Z+jet or γ+jet events, where a hadronic jet occurs back-to-back with either a Z-boson or a photon. The p_T response of Z-bosons and photos is very precisely measured in the detector, and can be compared to the hadronic jet in order to create a scaling factor applied as the absolute residual correction to be applied to the data.

5.2 Jet Energy Resolution

The measurement of the detector is not perfectly accurate. Jet Energy Resolution (JER) serves as a means to determine the accuracy of the detector's jet measurements. A comparison of GEN-level and RECO-level events from Monte Carlo simulation can be used to find the JER for the measurements in this analysis. RECO and GEN events are matched spatially, then the detector-reported p_T of a RECO jet is plotted over the p_T of the corresponding GEN jet. This process is repeated for every matched RECO-GEN jet pair in the given p_T bin, for all p_T bins. Ideally, the energy ratio should be as close to unity as possible, but in practice the ratios fall in a distribution around unity. The distribution in each bin is fitted with a Gaussian function, and the average ratio value per bin is plotted as a function of the p_T bin. An average close to one indicates a consistent response. The width of the Gaussian distribution describes the resolution of the response, with a more narrow distribution meaning a more accurate measurement. The Gaussian fits for p_T bins from 43 to 153 for $0.0 < |\eta| < 0.5$ AK3PF jets can be seen in Fig. 5.1. Plots containing the average (μ) and variance (σ) for $0.0 < |\eta| < 0.5$ AK3PF jets for p_T jet p_T bins can be

seen in Fig. A.1, along with a third plot showing the variance divided by the mean. The mean of both methods show a very close distribution around unity for the central η bin, deviating only by a maximum of 1%. The variances are tighter for the Gaussian function-derived entries, but both show expected tapering behavior correlating with higher p_T bin values. Finally, the polynomial fit of the variance to mean ratio can be used in the case of further comparison to NLO calculated simulation, a future area of interest but outside the scope of this analysis. JER histograms for the other η bins can be seen in the Appendix.

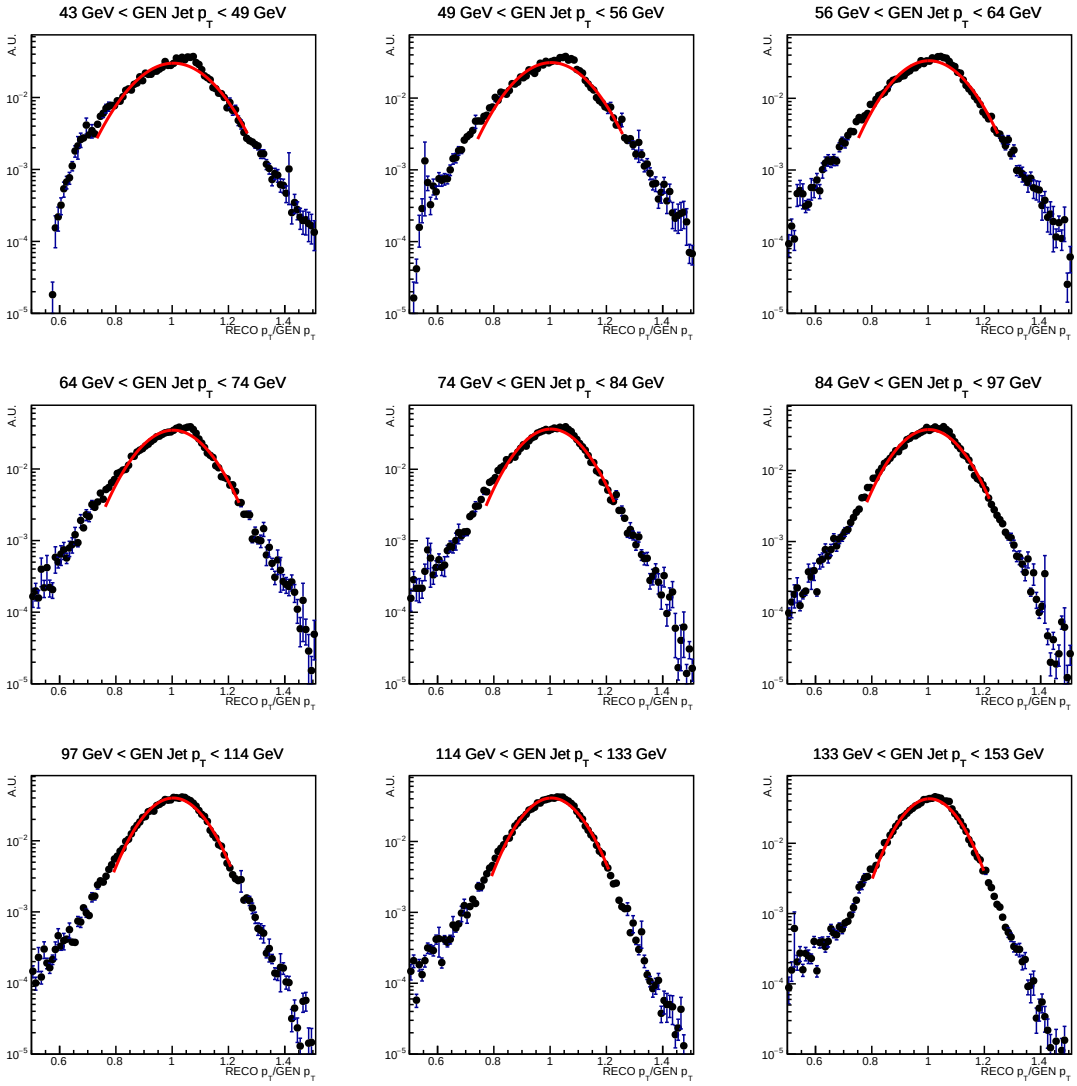

Figure 5.1: Black circles show RECO/GEN energy distributions for $0.0 < |\eta| < 0.5$ AK3PF jets for p_T bins from 43 GeV to 153 GeV. The red line indicates a Gaussian function fit to the distributions.

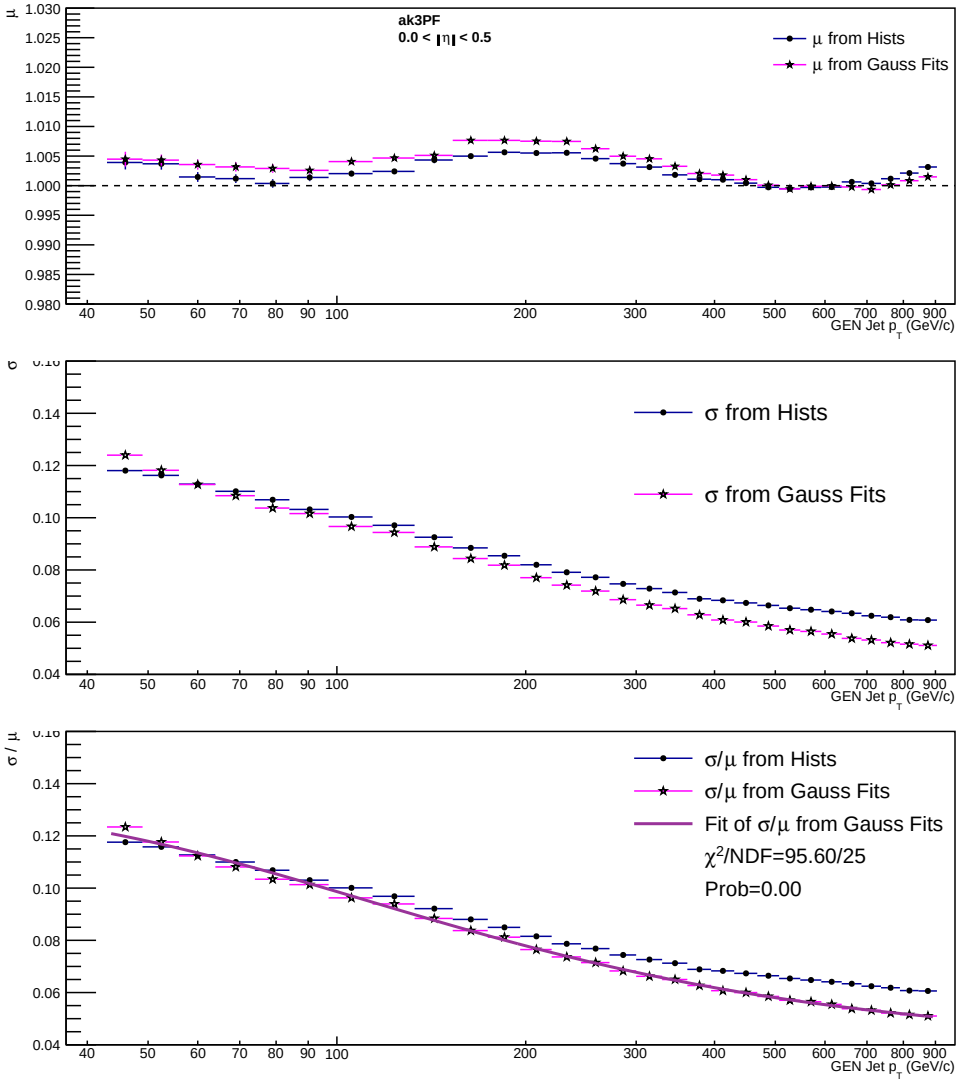

Figure 5.2: TOP: The RECO/GEN energy ratio mean per p_T bin derived from the histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). The dashed line indicates unity. MIDDLE: Variance of RECO/GEN energy ratio per p_T bin derived from the histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). BOTTOM: Variance to mean ratio for values derived from histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). The dark purple line indicates the fit function of the Gaussian-derived variance to mean ratio.

Chapter 6: Inclusive Jet Measurements

6.1 Unfolding

Although the JEC applied to the data in order to reach the RECO stage can correct for many of the detector effects, it cannot correct for the bin-to-bin p_T migrations that occur during the process. While the JEC corrects based on the average p_T of a bin, there will always be statistical variance within the bin resulting in several jets being over- and under-corrected per bin. In order to correct for these migrations, the d'Agostini "Bayesian" unfolding technique is used to infer the "true" distribution from the distorted one recorded by the detector [33]. Using Monte Carlo simulation and the GEANT4 detector simulation, a response matrix can be made correlating the RECO MC back to its GEN value while also taking into account statistical uncertainties. In Fig. 6.1 the response matrix is shown for all bins of η. The response matrix is a map of GEN jet p_T to RECO jet p_T based on Monte Carlo simulation with PYTHIA8. Entries on the direct diagonal show an exact match between GEN and RECO jet p_T, while entries that stray from the diagonal line are caused by detector resolution effects. Using an approach based on Bayes' Theorem, an unfolding matrix is created in order to unfold the measured data.

In order to confirm the validity of the unfolding algorithm, a closure test is run, which uses the MC simulation as a tool to unfold itself. The PYTHIA8 MC simulation is

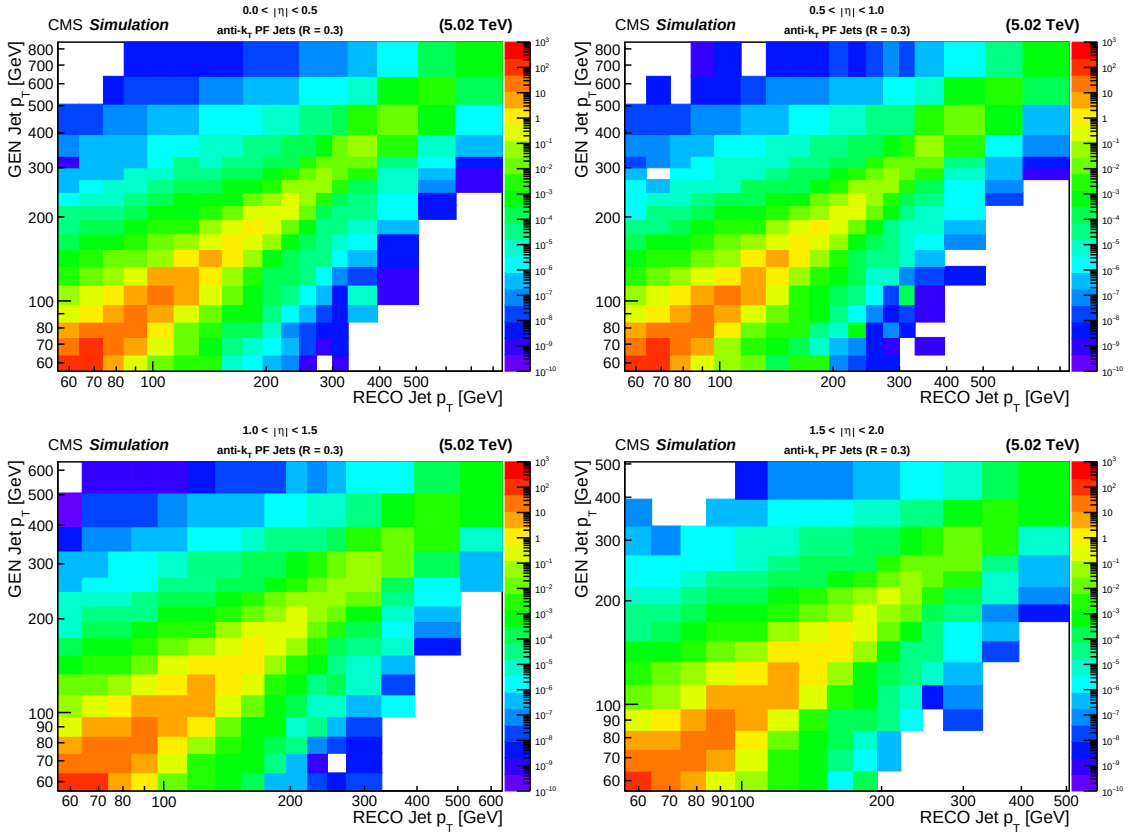

Figure 6.1: Response matrix for all η bins of AK3PF jets, showing map of matching RECO and GEN jet p_T from PYTHIA8 Monte Carlo simulation.

arbitrarily divided into two equal-sized portions, labeled as "test" and as "truth". The truth portion is fed into the GEANT 4 simulated detector in order to simulate detector effects, effectively creating a simulated version of the RECO-level data. This can be compared to the original state of the truth sample to create a response matrix. The process of feeding the sample into GEANT to create a simulation with detector effects is repeated for the test sample. An unfolding matrix created from the truth sample is then used to unfold the test sample, and the results compared. The unfolding matrix is used to unfold the truth sample as well to provide a very basic sanity check of the procedure. The resulting ratio is expected to be at a value of 1 with standard statistical fluctuations between the two

sample sets. The closure test for $0.0 < |\eta| < 0.5$ AK3PF jets used in this analysis can be seen in Fig. B.1. The full set of plots for all rapidity bins is available in the appendix. The unfolded test PYTHIA8 follows expected behavior, showing only statistical fluctuations around unity. RECO test and truth are also shown, at a ratio of around 1.1. As these represent the post-unfolding events, it can be inferred that the rise in the ratio value is due to detector effects.

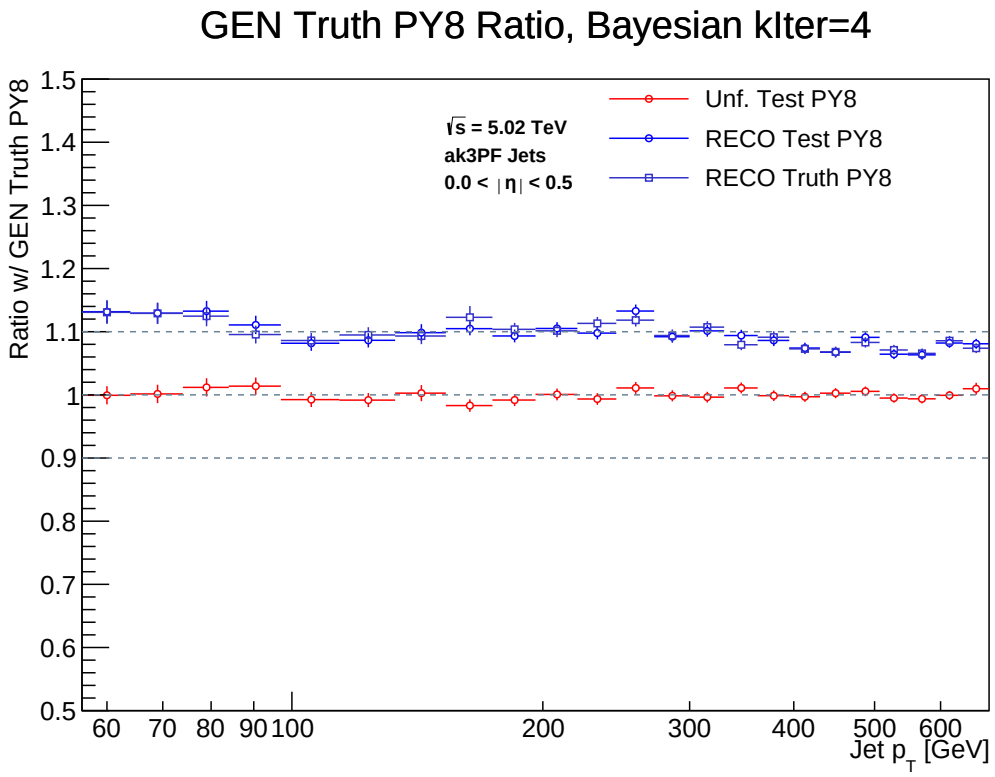

Figure 6.2: Closure test for $0.0 < |\eta| < 0.5$ AK3PF jets. Red circles show the unfolded Test PY8 simulation as a ratio to the GEN PY8 sample. Blue circles show the RECO-level Test sample, displaying a raised ratio to GEN Truth. Blue squares show RECO-level Truth, also showing a raised ratio. The increase can be attributed to detector effects. The unfolded Test PY8 shows good accordance with the ideal value of 1. Dashed lines represent $\pm 10\%$ from unity. The full set of plots for all rapidity bins is available in the appendix.

The folding ratios for the analysis can be seen in Fig C.4. These ratios are an inverse of the unfolding process and another check on the validity of the unfolding algorithm.

Instead of unfolding the RECO to compare to the GEN-level truth, folding ratios compare the folded GEN to the RECO-level. As before, a value around 1 with statistical fluctuations is expected. In this context, real jets that have been reconstructed with a highly different p_T from the GEN level are referred to as "fakes". These jets are automatically subtracted from the sample during the unfolding process by the rooUnfold code; however, if the sample is re-folded, the fakes must be manually added back in. The results shown here for $0.0 < |\eta| < 0.5$ AK3PF jets follow the expected pattern, aligning at unity with small statistical fluctuations. Results for other η bins can be found in the appendix.

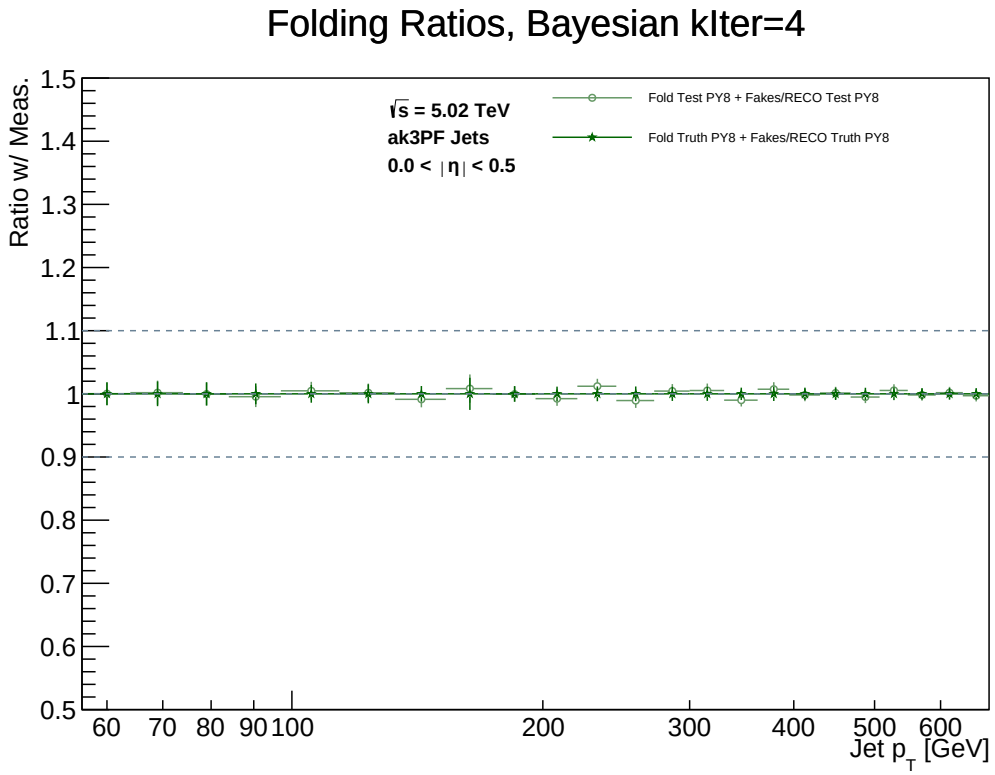

Figure 6.3: Folding Ratios for $0.0 < |\eta| < 0.5$ AK3PF jets. Folded GEN-level events are shown as a ratio to RECO-level PY8 Monte Carlo simulation. Dashed lines represent $\pm 10\%$ from unity. The full set of plots for all rapidity bins is available in the appendix.

6.1.1 Iterative Ratio

The χ^2 difference between different iterations of the D'Agostini unfolding algorithm can be seen in Fig. 6.4. Once the value of χ^2 falls below 0.01, there is no significant change in further iterations outside of statistical fluctuations. The values shown here are unusually low in comparison to similar jet studies at other radii, and merit further investigation.

6.1.2 Correction Factor

The jet data cross section spectra for 5 TeV AK3PF jets are shown in Fig. 6.5. The cross sections derived from RECO-level data and those derived from unfolded data are plotted as a function of jet p_T in order to demonstrate the effect of unfolding on the measurement. The outermost bins show the largest difference between the measured and unfolded data, likely due to the combination of low statistics and increased detector effects inherent to these regions. In Fig. 6.6, the unfolded data is plotted as a ratio of RECO-level data, demonstrating the unfolding correction factor as a ratio. A relatively flat correction factor indicates that the unfolding results are stable.

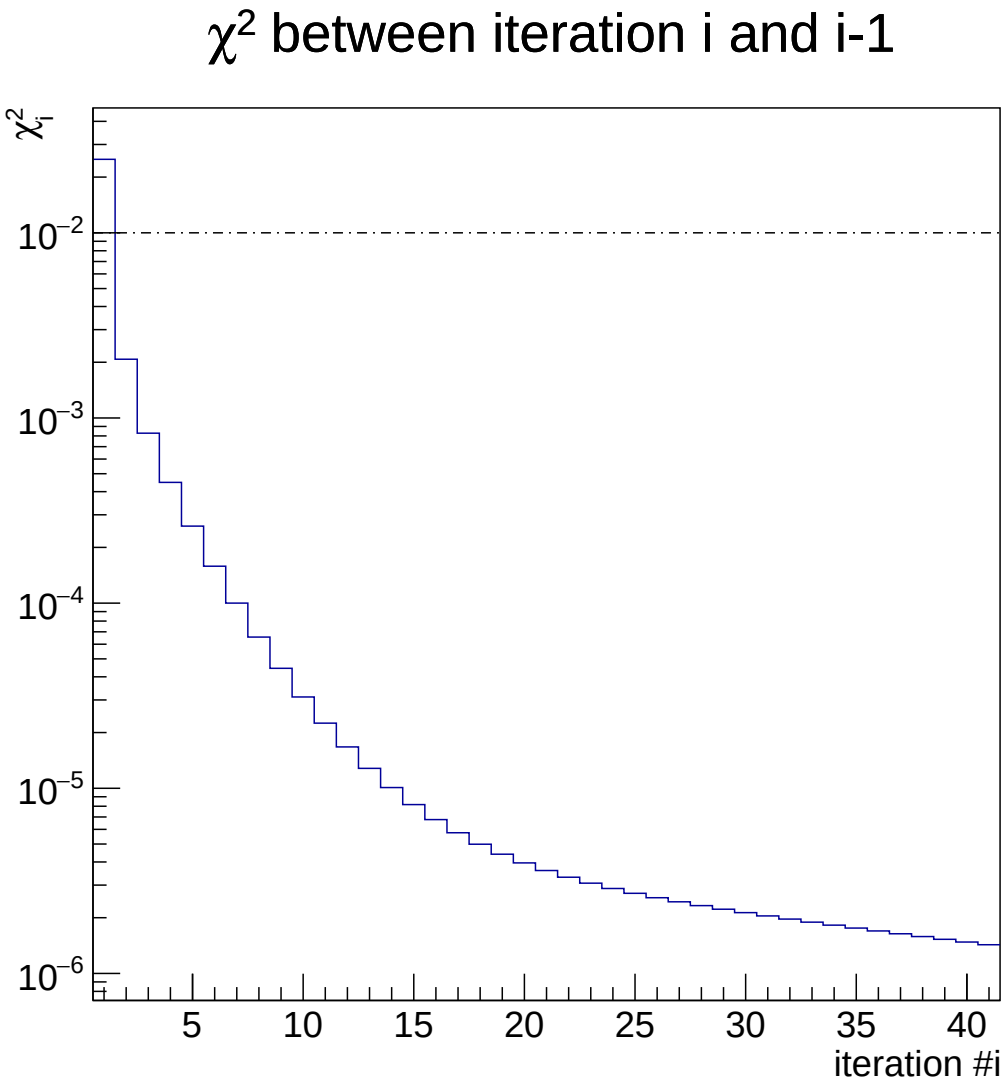

Figure 6.4: Iterative Ratio using the D'Agostini Bayes-based iterative unfolding algorithm.

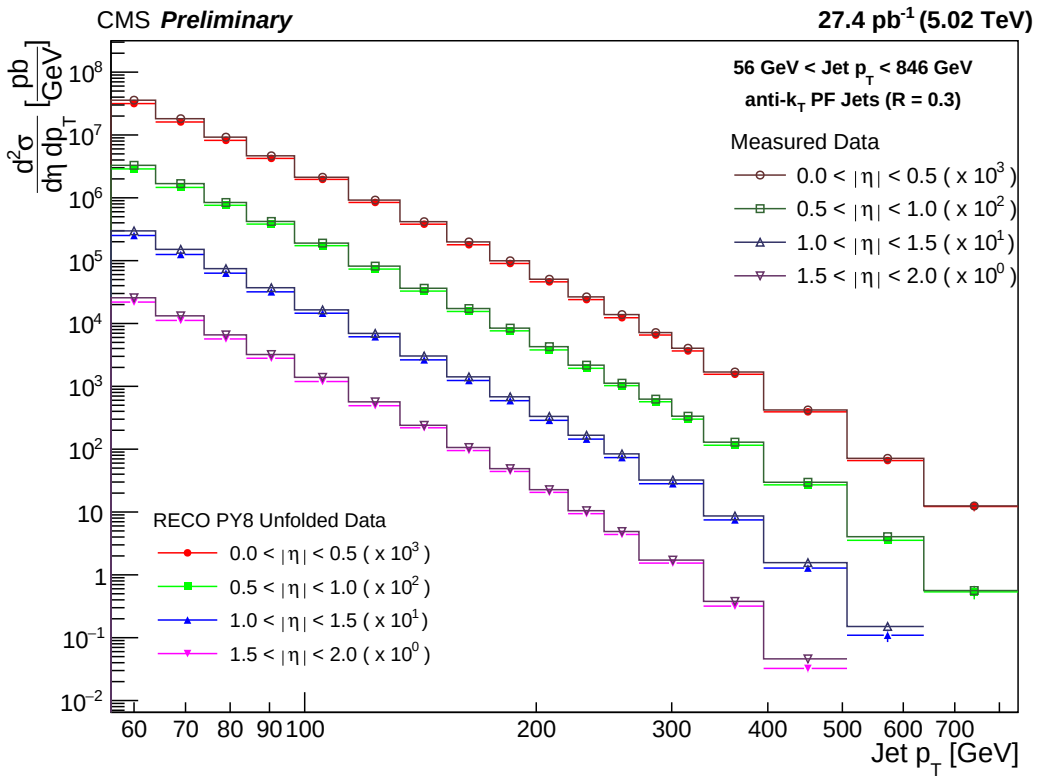

Figure 6.5: The jet data spectra for AK3PF jets, with η bins scaled for clearer presentation.

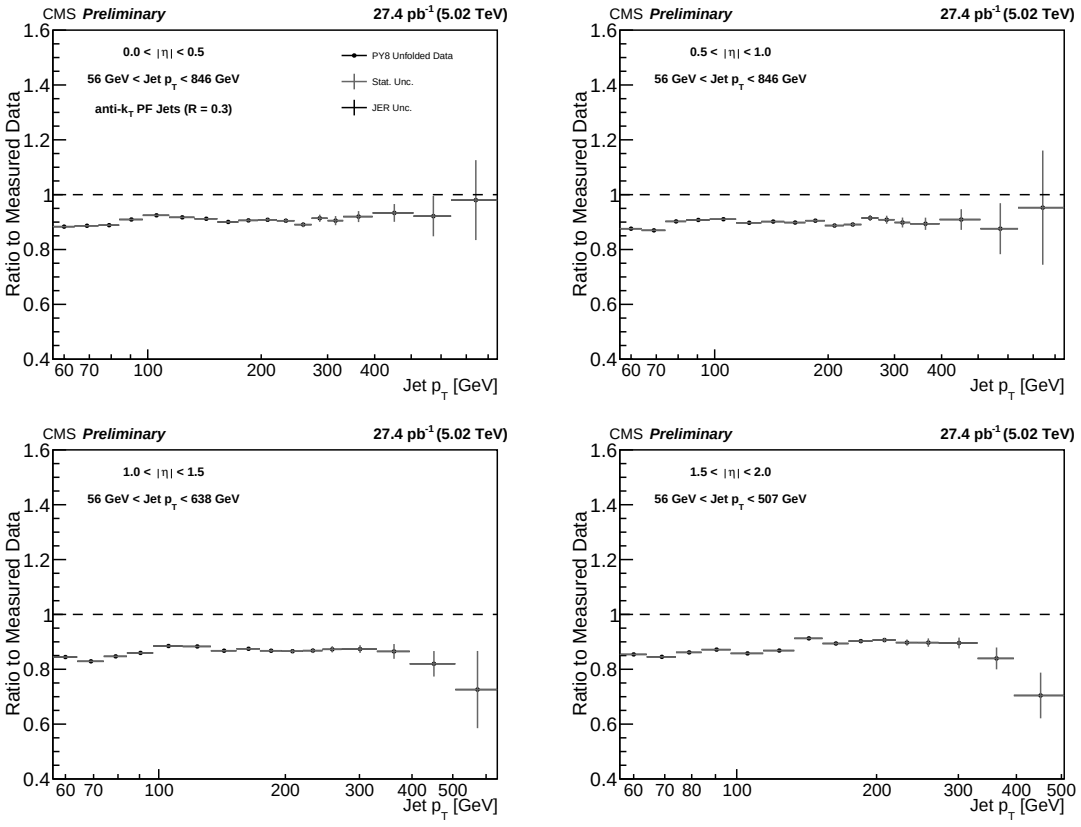

Figure 6.6: Correction factor for AK3PF jets in all four η bins.

Chapter 7: Results

7.1 Comparison with PYTHIA8 Monte Carlo Simulation

Shown in 7.1 are the PYTHIA8-unfolded cross sections for AK3PF jets in bins of $|\eta|$ from 0 to 2.0, compared with PYTHIA8 GEN-level simulation. Cross sections are scaled for readability.

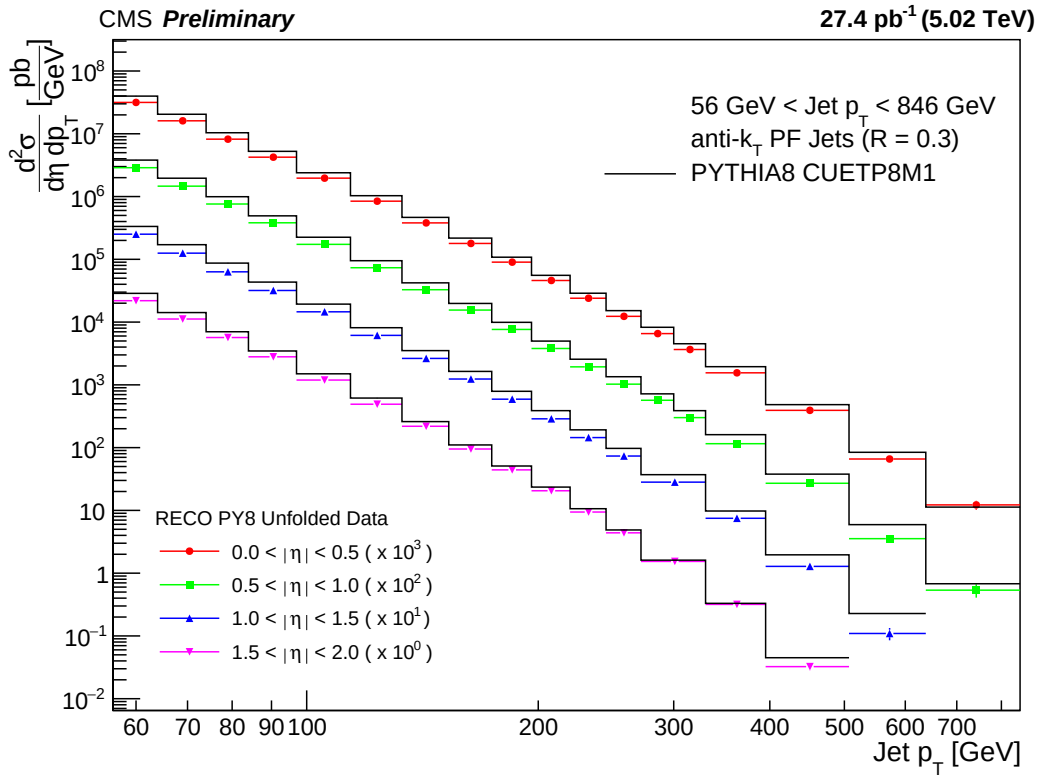

Figure 7.1: PYTHIA8-unfolded cross sections for $0.0 < |\eta| < 2.0$ jets, compared with PYTHIA8 GEN-level simulation. Cross sections are scaled for greater readability.

The ratio of the PY8-unfolded data to PYTHIA8 MC simulation is shown for bins of $|\eta|$ from 0 to 2.0 in figure 7.2. The gray bars indicate the combined statistical uncertainty of the data and MC simulation, the red lines show the JEC systematic uncertainty, and the black bars show the combined total uncertainty.

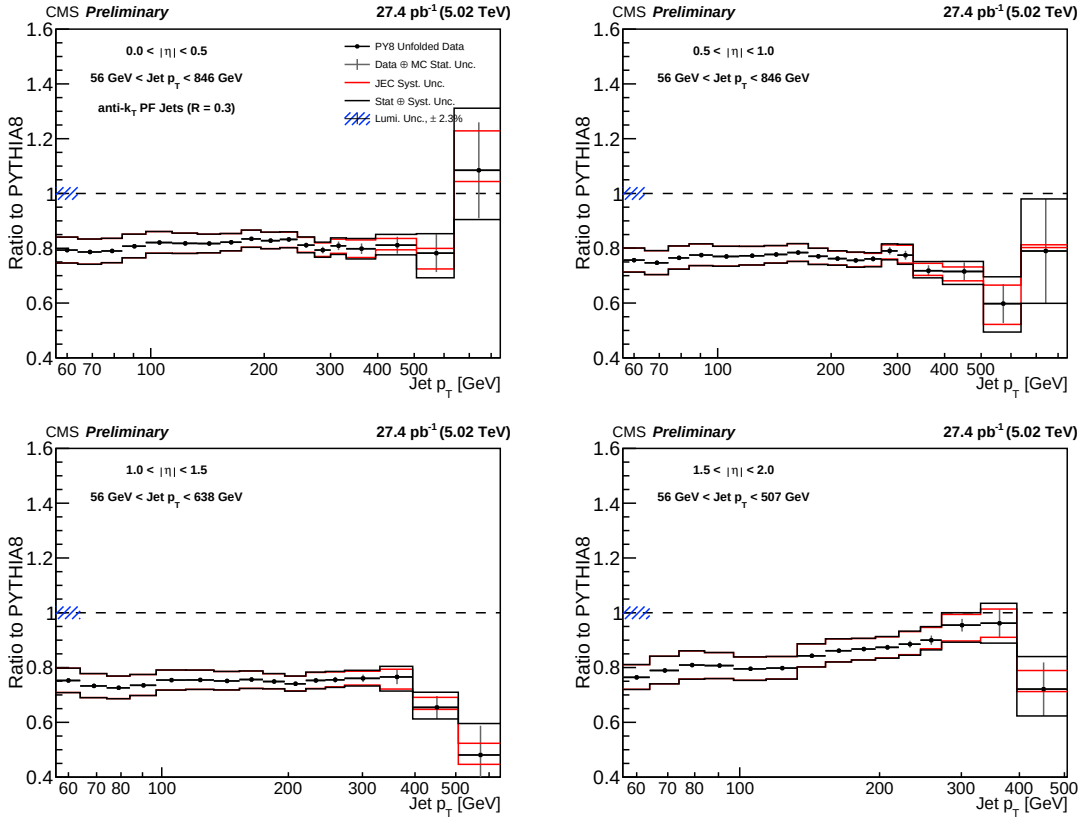

Figure 7.2: Ratio of PYTHIA8 unfolded data to PYTHIA8 GEN-level MC simulation for $0.0 < |\eta| < 2.0$ AK3PF jets. Combined data and MC systematic uncertainty (gray), JEC systemic uncertainty (red), and total systemic uncertainty (black) are also shown.

The ratio of the measured data to the MC simulation sits at approximately 80% for all η bins, with deviations mainly above 400 GeV p_T and in the outermost η bin. The deviation in the outermost η bin is not entirely unsurprising due to a combination of lower statistics and more pronounced detector effects in this area of the detector. Although not at unity, the consistency of the data/MC ratio at roughly 80% for the middle and lower p_T

bins is a reassuring sign. Given the constraints of the error distributions, the discrepancy between simulation and data is unlikely to have been caused by missteps in the analysis procedure. The behavior of the ratio at the highest p_T bins, rising sharply in the innermost η bins, dropping in the third bin, and rising then dropping in the outermost η bin, is not ideal but matches the behavior of the data/PY8 ratio in preliminary results from an R = 0.4 analysis carried out on the same dataset and PY8 Monte Carlo simulation, seen in Fig. 7.3 [6]. Investigation with more advanced simulation at next-to-leading order (NLO) and next-to-next-to leading order (NNLO) simulations is merited in order to gain further insight.

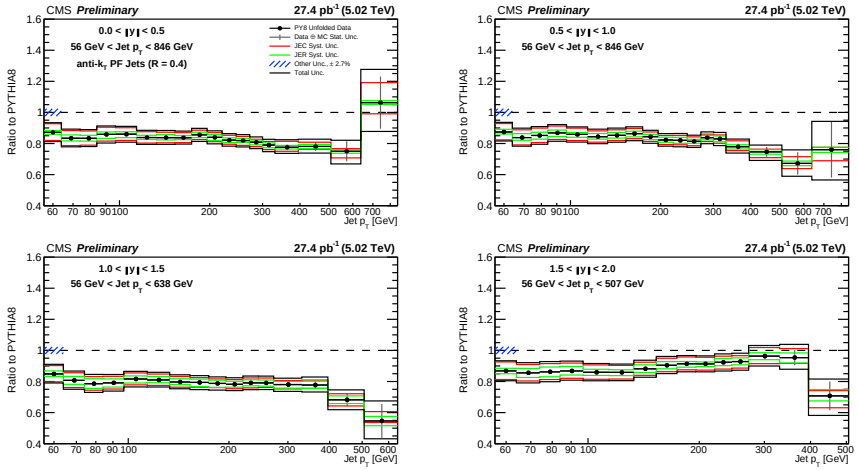

Figure 7.3: Ratio of PYTHIA8 unfolded data to PYTHIA8 GEN-level MC simulation for $0.0 < |\eta| < 2.0$ AK4PF jets. Combined data and MC systematic uncertainty (gray), JEC systemic uncertainty (red), JER uncertainty (green) and total systemic uncertainty (black) are also shown. Taken from [6].

7.2 Conclusion

The double-differential cross sections for AK3PF jets at $\sqrt{s} = 5.02$ TeV in terms of pseudorapidity (η) and transverse momentum (p_T) have been presented. The raw detector

response was corrected by applying Jet Energy Corrections for absolute and relative detector effects in p_T and η. Unfolding to correct for detector effects in terms of p_T bin migration was carried out using the PY8 MC LO-derived folding matrix. Comparison of data and PY8 MC simulation shows that LO theory consistently predicts the cross sections at around 80% of the values found in data, but begins to differ sharply at higher p_T bins and, to a lesser extent, in more outer η regions of the detector.

Appendix A: Jet Energy Resolution

The following figures are the Jet Energy Resolution histograms for all η bins.

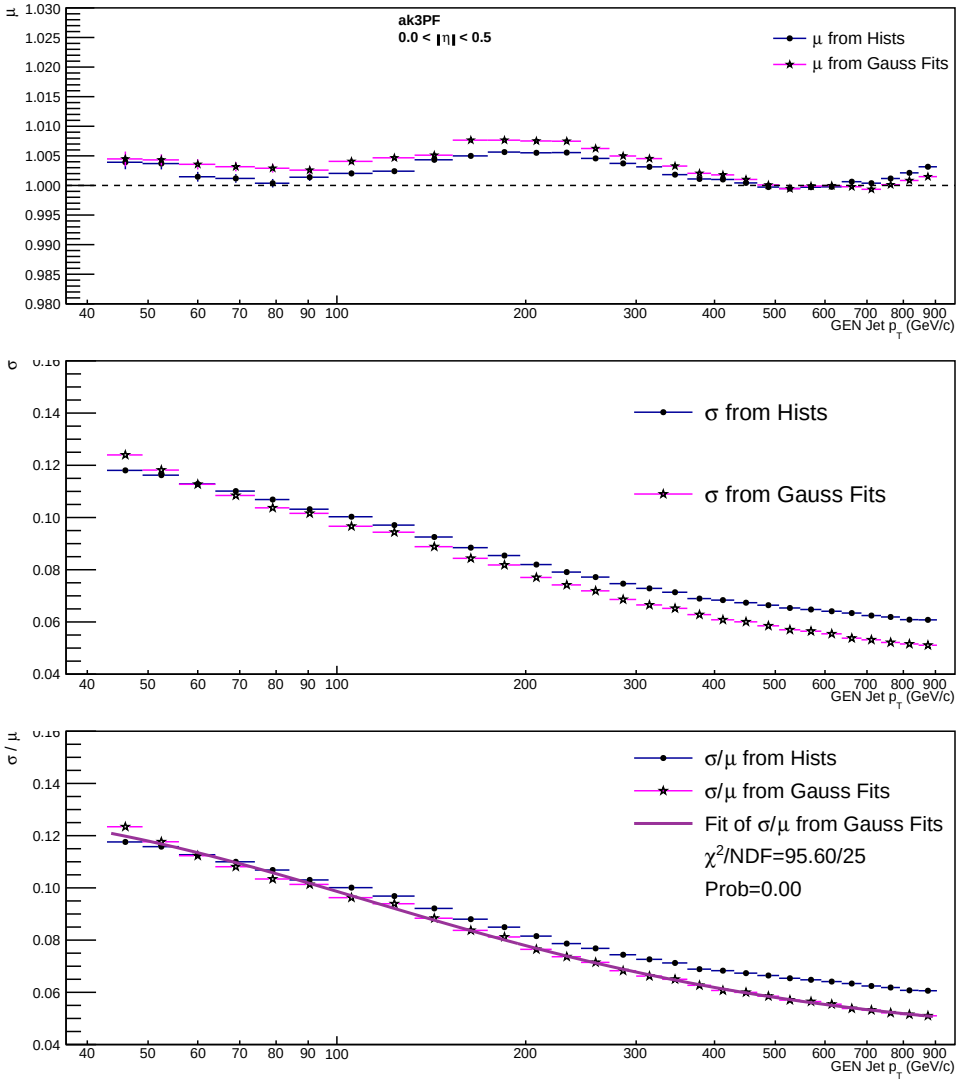

Figure A.1: JERs histograms for $0.0 < \eta < 0.5$. TOP: The RECO/GEN energy ratio mean per p_T bin derived from the histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). Dashed line indicates unity. MIDDLE: Variance of RECO/GEN energy ratio per p_T bin derived from the histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). BOTTOM: Variance to mean ratio for values derived from histogram entries (black circles on blue) and the Gaussian fits of the histograms (black stars on magenta). The dark purple line indicates the fit function of the Gaussian-derived variance to mean ratio.

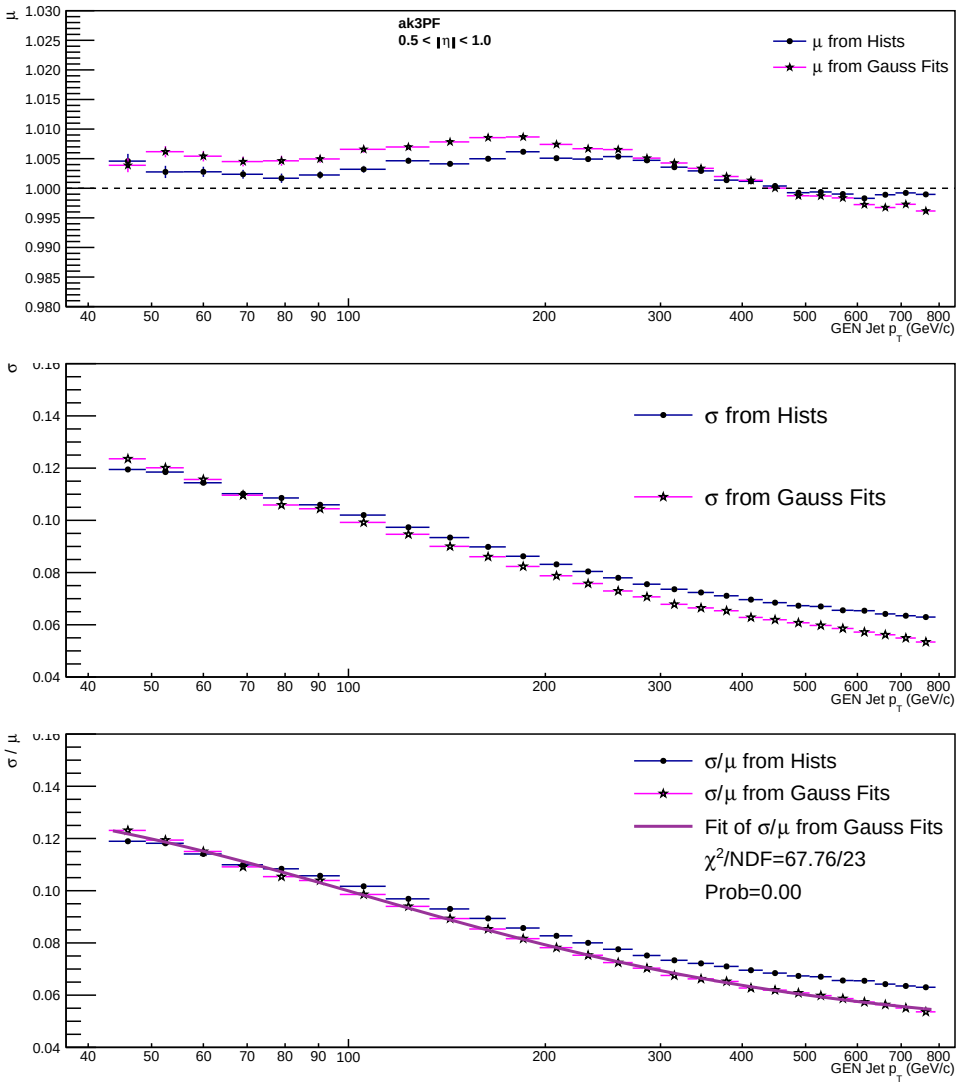

Figure A.2: JERs histograms for $0.5 < \eta < 1.0$.

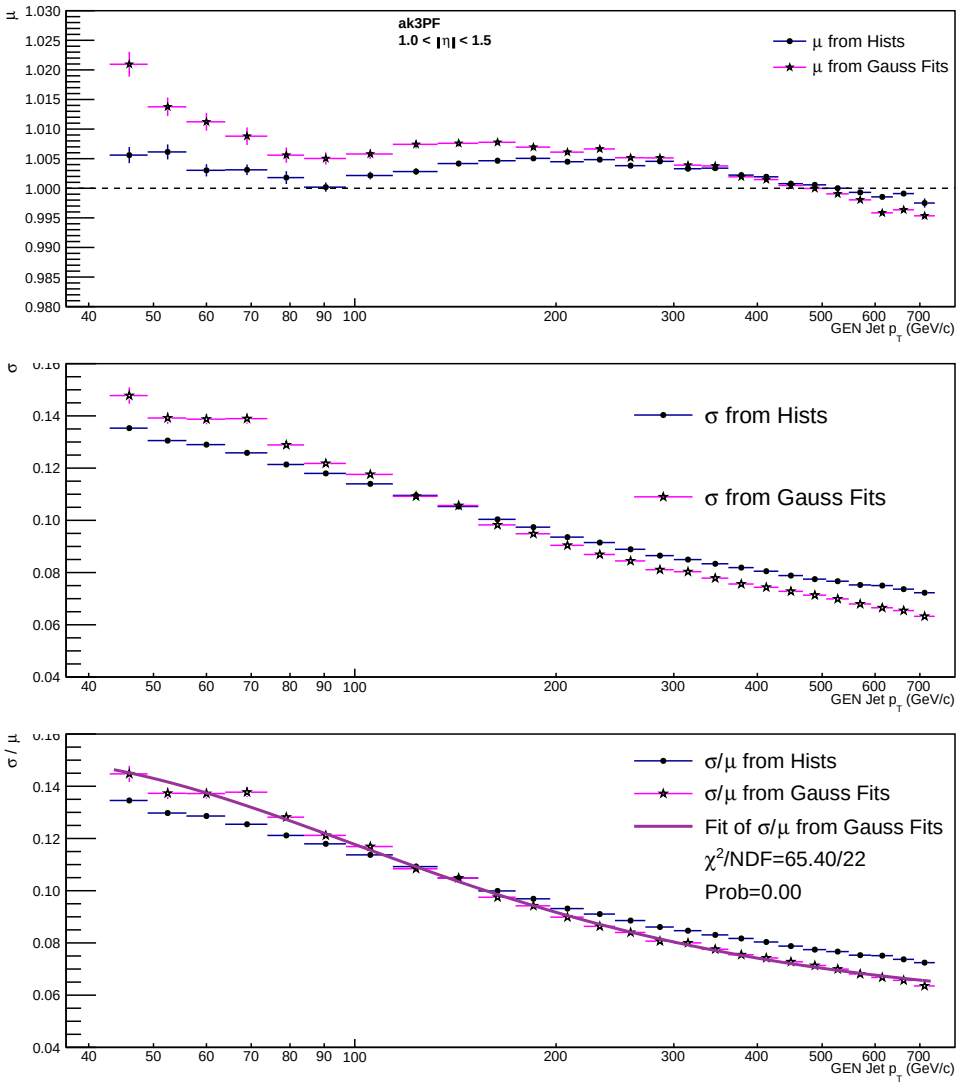

Figure A.3: JERs histograms for $1.0 < \eta < 1.5$.

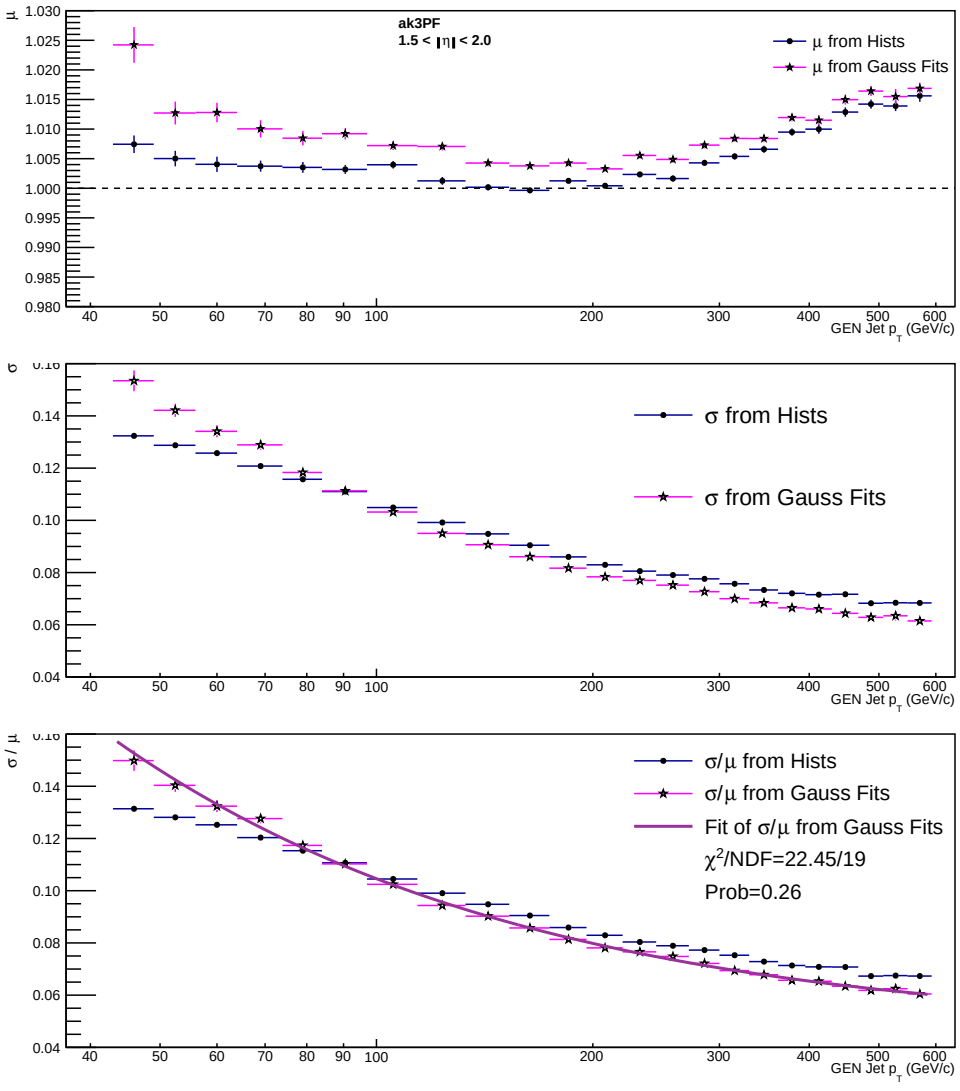

Figure A.4: JERs histograms for $1.5 < \eta < 2.0$.

Appendix B: Closure Tests

The following figures are the closure test histograms for all η bins.

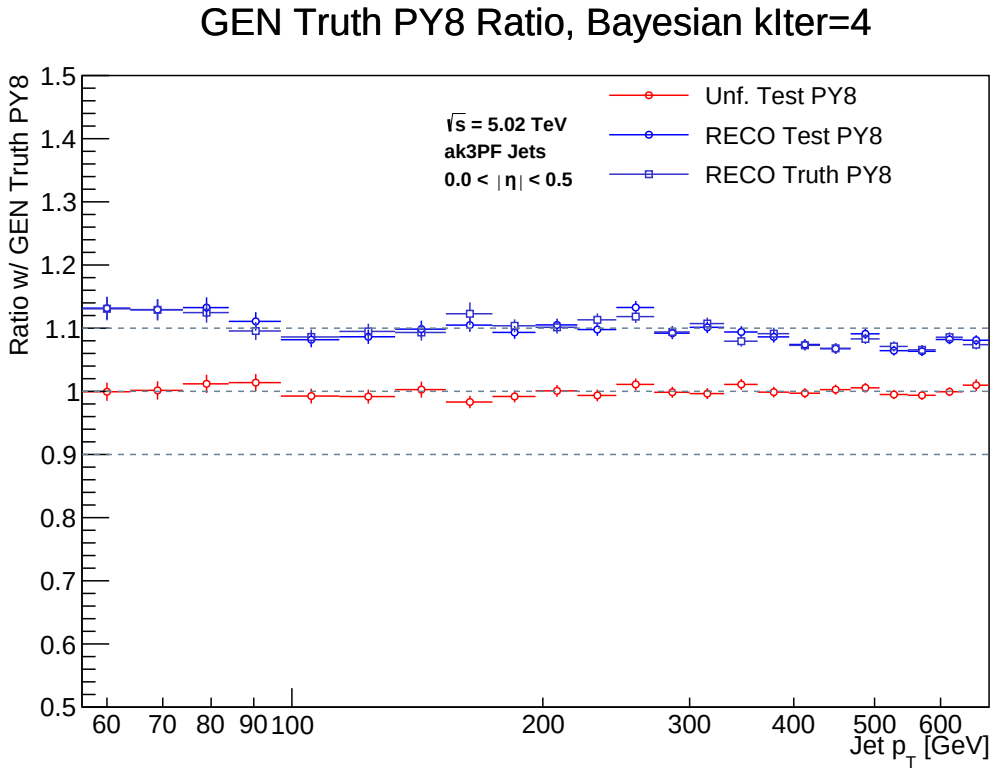

Figure B.1: Closure test for $0.0 < |\eta| < 0.5$ AK3PF jets. Red circles show the unfolded Test PY8 simulation as a ratio to the GEN PY8 sample. Blue circles show the RECO-level Test sample, displaying a raised ratio to GEN Truth. Blue squares show RECO-level Truth, also showing a raised ratio. The increase can be attributed to detector effects. The unfolded Test PY8 shows good accordance with the ideal value of 1. Dashed lines represent $\pm 10\%$ from unity.

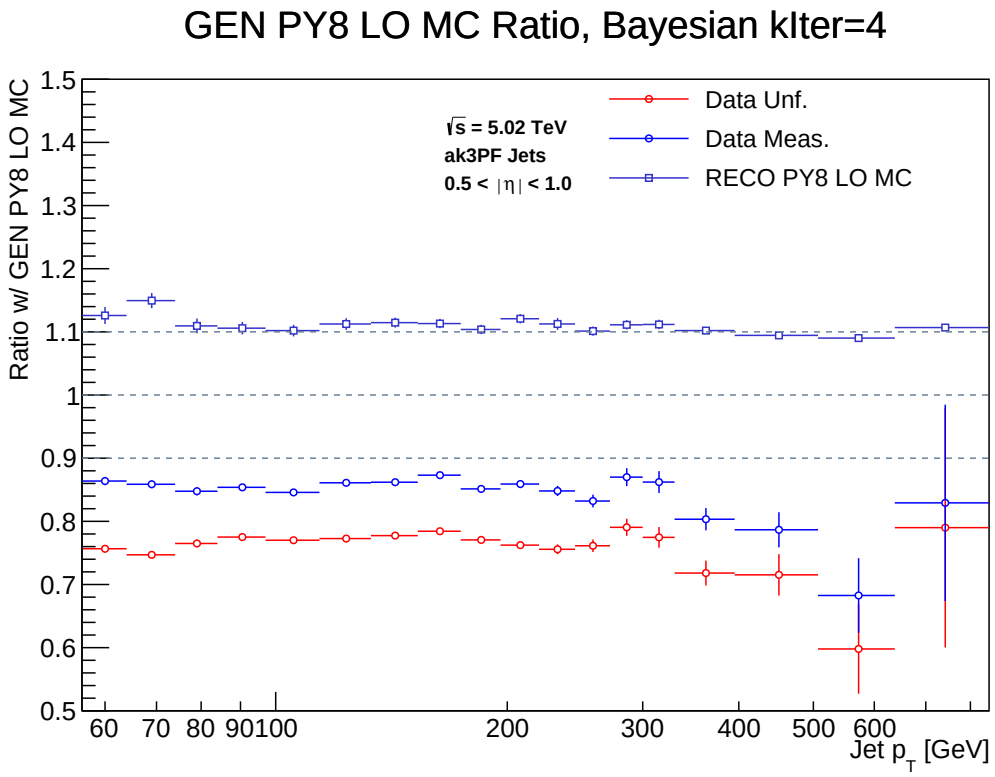

Figure B.2: Closure test for $0.5 < |\eta| < 1.0$ AK3PF jets.

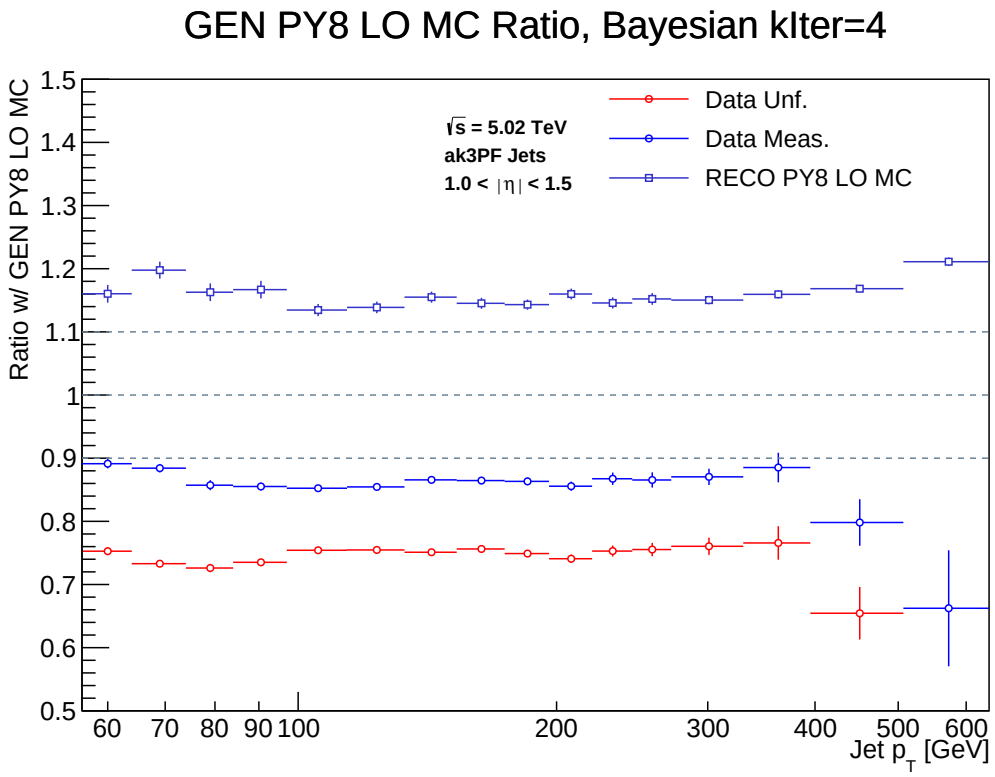

Figure B.3: Closure test for $1.0 < |\eta| < 1.5$ AK3PF jets.

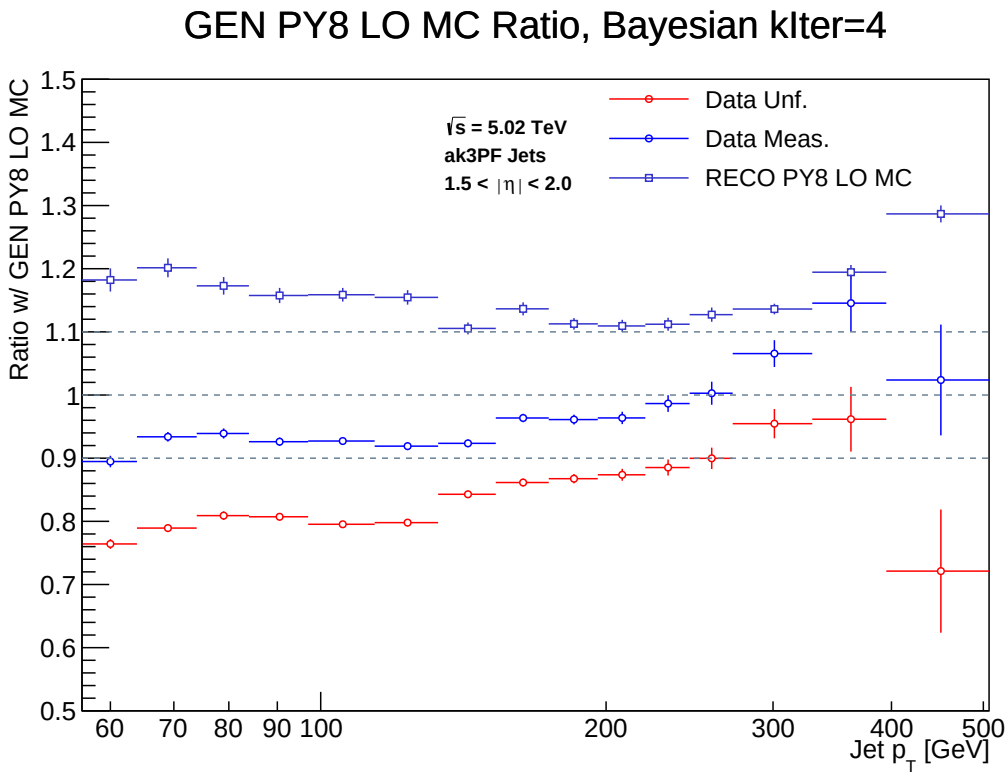

Figure B.4: Closure test for $1.5 < |\eta| < 2.0$ AK3PF jets.

Appendix C: Folding Ratios

The following figures are the folding ratio histograms for all η bins.

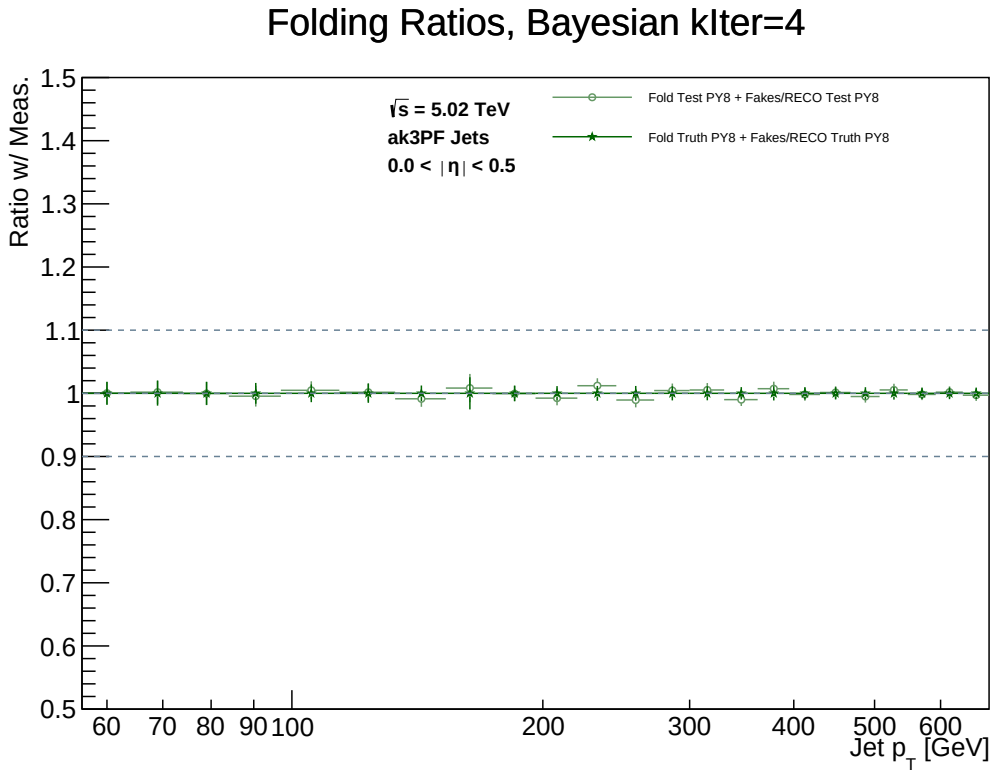

Figure C.1: Folding Ratios for $0.0 < |\eta| < 0.5$ AK3PF jets. Folded GEN-level events are shown as a ratio to RECO-level PY8 Monte Carlo simulation. Dashed lines represent $\pm 10\%$ from unity.

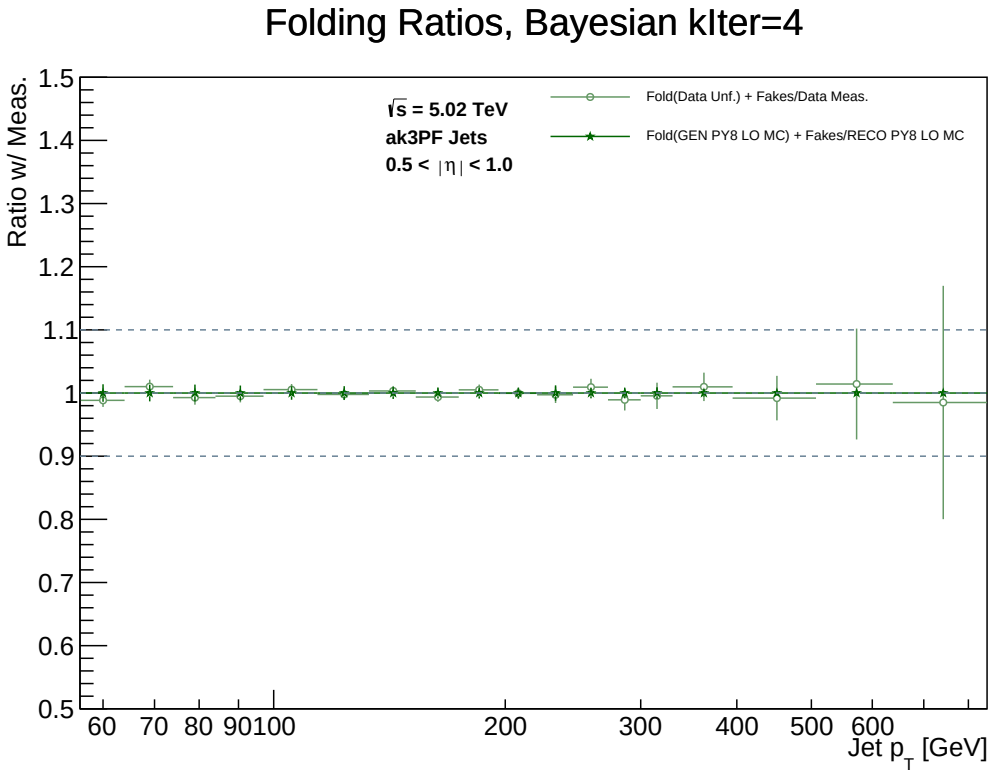

Figure C.2: Folding Ratios for $0.5 < |\eta| < 1.0$ AK3PF jets.

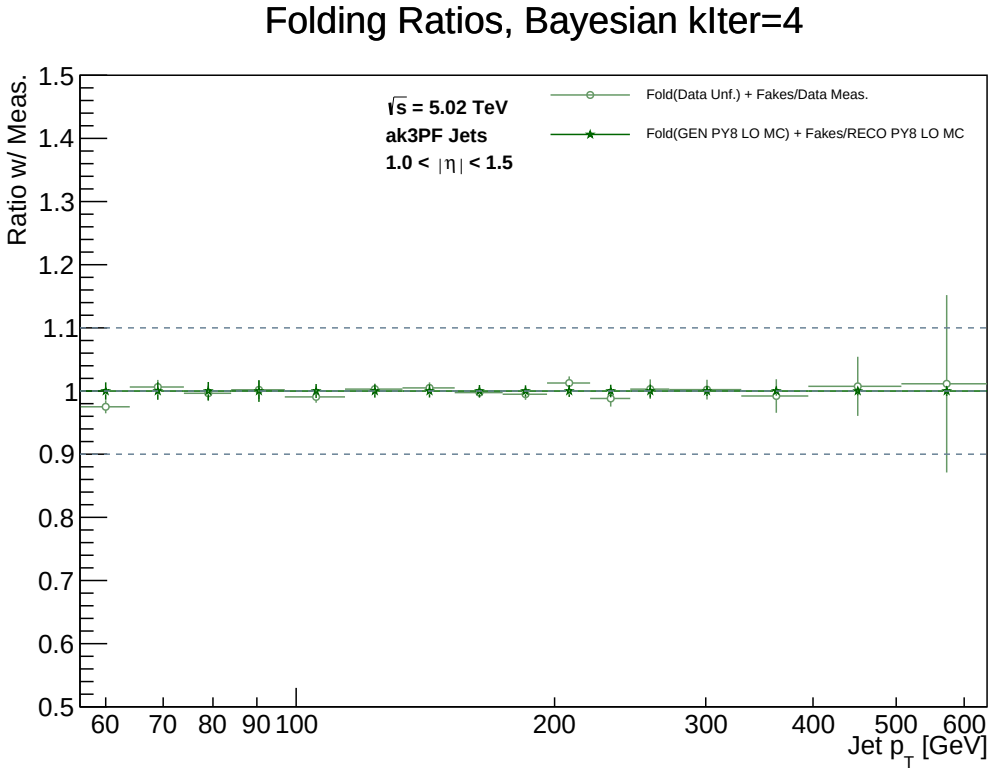

Figure C.3: Folding Ratios for $1.0 < |\eta| < 1.5$ AK3PF jets.

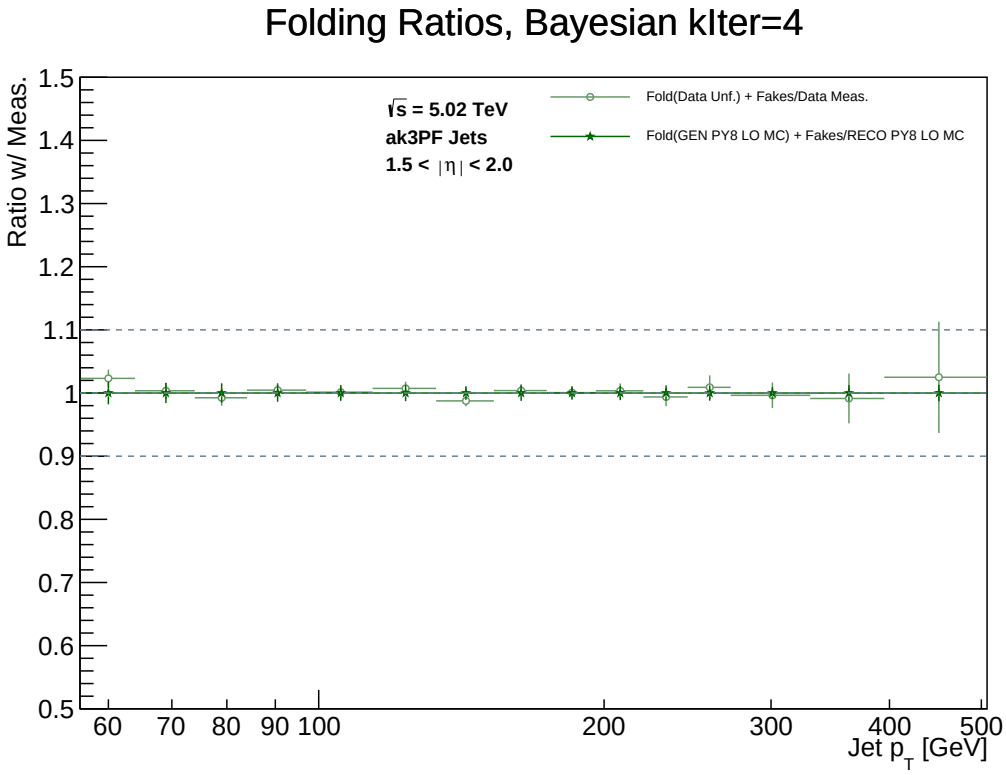

Figure C.4: Folding Ratios for $1.5 < |\eta| < 2.0$ AK3PF jets.

Bibliography

[1] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, and et al. Search for supersymmetry at the lhc in events with jets and missing transverse energy. *Physical Review Letters*, 107(22), Nov 2011.

[2] V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, R. Frühwirth, and et al. Search for microscopic black hole signatures at the large hadron collider. *Physics Letters B*, 697(5):434–453, Mar 2011.

[3] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[4] Christof Roland. Heavy-ion physics at the LHC with the compact muon solenoid detector. In *39th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions*, 5 2004.

[5] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.

[6] Inclusive Jet Cross Section for R = 0.4 Anti-k_T Particle Flow Jets using $\sqrt{S} = 5.02$ TeV pp Collisions with the CMS Detector. Technical report, CERN, Geneva, 2016.

[7] Bogdan Povh, Klaus Rith, Christoph Sholz, Frank Zetsche, and Martin Lavelle. *Particles and Nuclei: An Introduction to the Physical Concepts*. Graduate Texts in Physics. Springer-Verlag, Berlin Heidelberg, 2008.

[8] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[9] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012.

[10] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[11] G. L. Bayatian et al. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.

[12] S. Chatrchyan et al. Measurement of the rapidity and transverse momentum distributions of z bosons in pp collisions at sqrt(s)=7 tev. *Phys. Rev. D*, 85:032002, Feb 2012.

[13] Ying Liu. private communication.

[14] Paolo Azzurri. The CMS Silicon Strip Tracker. *J. Phys.: Conf. Ser.*, 41:127–134. 8 p, Dec 2005. 8 pages, 8 figures, talk given at XIX EPS NPDC Conference on New Trends in Nuclear Physics Applications and Technology, September 5-9, 2005 Pavia, Italy Subj-class: Instrumentation and Detectors.

[15] CMS Collaboration. Commissioning and performance of the cms silicon strip tracker with cosmic ray muons. *Journal of Instrumentation*, 5(03):T03008–T03008, Mar 2010.

[16] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[17] CMS Collaboration. Performance of the cms hadron calorimeter with cosmic ray muons and lhc beam data. *Journal of Instrumentation*, 5(03):T03012–T03012, Mar 2010.

[18] Abdullin S. et al. Design, performance, and calibration of the cms hadron-outer calorimeter. *Eur. Phys. J. C*, 57(3):653–663, 2008.

[19] Domenico Campi et al. Commissioning of the CMS Magnet. *IEEE Trans. Appl. Supercond.*, 17(2):1185–1190, 2007.

[20] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997.

[21] Valentina Gori. The cms high level trigger. *International Journal of Modern Physics: Conference Series*, 31:1460297, Jan 2014.

[22] C. Foudas. The cms level-1 trigger at lhc and super-lhc, 2008.

[23] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, Jun 2015.

[24] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[25] R Frühwirth, P Kubinec, Winfried A Mitaroff, and Meinhard Regler. Vertex reconstruction and track bundling at the LEP collider using robust algorithms. *Comput. Phys. Commun.*, 96:189–208. 23 p, Mar 1996.

[26] W Erdmann. Vertex reconstruction at the CMS experiment. *Journal of Physics: Conference Series*, 110(9):092009, may 2008.

[27] Florian Beaudette. The CMS Particle Flow Algorithm. In *International Conference on Calorimetry for the High Energy Frontier*, 2013.

[28] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[29] Pileup Removal Algorithms. Technical report, CERN, Geneva, 2014.

[30] A. Santocchia. Jet energy correction in cms. In *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, pages 2299–2307, 2009.

[31] Plans for Jet Energy Corrections at CMS. Technical report, CERN, Geneva, Jul 2008.

[32] V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al. Jet energy scale and

resolution in the cms experiment in pp collisions at 8 tev. *Journal of Instrumentation*, 12(02):P02014–P02014, Feb 2017.

[33] Giulio D'Agostini. *Bayesian reasoning in high-energy physics: principles and applications*. CERN Yellow Reports: Monographs. CERN, Geneva, 1999.